

Tesis de Posgrado

Organización y expresión genómica en Trypanosoma cruzi : Proteínas ribosomales P y secuencias repetidas

Vazquez, Martín Pablo

1995

Tesis presentada para obtener el grado de Doctor en Ciencias
Biológicas de la Universidad de Buenos Aires

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Vazquez, Martín Pablo. (1995). Organización y expresión genómica en Trypanosoma cruzi : Proteínas ribosomales P y secuencias repetidas. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_2796_Vazquez.pdf

Cita tipo Chicago:

Vazquez, Martín Pablo. "Organización y expresión genómica en Trypanosoma cruzi : Proteínas ribosomales P y secuencias repetidas". Tesis de Doctor. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 1995.

http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_2796_Vazquez.pdf

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires

**UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES**

**ORGANIZACIÓN Y EXPRESIÓN GENÓMICA EN *Trypanosoma cruzi*:
PROTEÍNAS RIBOSOMALES P Y SECUENCIAS
REPETIDAS.**

MARTIN PABLO VAZQUEZ

**TESIS PARA OPTAR POR EL TÍTULO DE DOCTOR EN
CIENCIAS BIOLÓGICAS.**

DIRECTOR: Dr. MARIANO JORGE LEVIN

1995

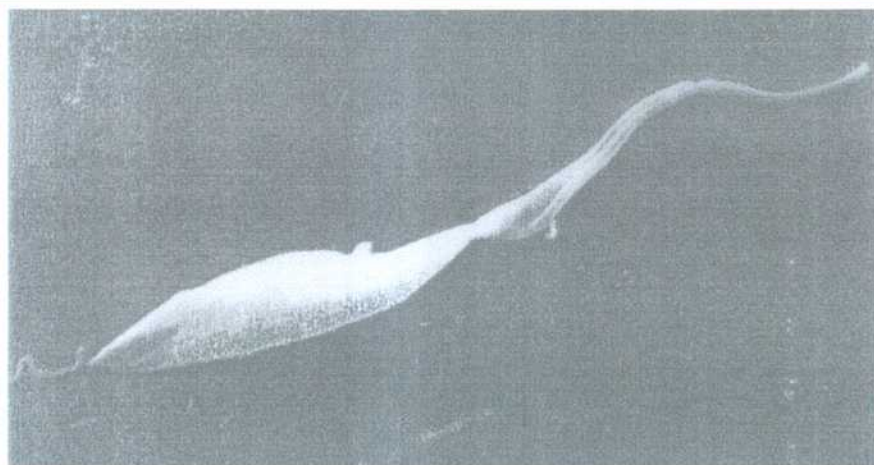
**INSTITUTO DE INVESTIGACIONES EN INGENIERIA GENÉTICA Y
BIOLOGÍA MOLECULAR (INGEBI)**

TESIS DE DOCTORADO

ORGANIZACION Y EXPRESION

GENOMICA EN *Trypanosoma*

***cruzi*: proteínas ribosomales P
y secuencias repetidas.**



Lic. MARTIN PABLO VAZQUEZ

**TESIS PARA OPTAR AL TITULO DE DOCTOR EN
CIENCIAS BIOLOGICAS
UNIVERSIDAD DE BUENOS AIRES.
FACULTAD DE CIENCIAS EXACTAS Y
NATURALES.**

**DIRECTOR: Dr. MARIANO JORGE LEVIN
1995.**

**UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES**

**ORGANIZACIÓN Y EXPRESIÓN GENÓMICA EN *Trypanosoma cruzi*:
PROTEÍNAS RIBOSOMALES P Y SECUENCIAS
REPETIDAS.**

MARTIN PABLO VAZQUEZ

**TESIS PARA OPTAR POR EL TÍTULO DE DOCTOR EN
CIENCIAS BIOLÓGICAS.**

DIRECTOR: Dr. MARIANO JORGE LEVIN

1995

**INSTITUTO DE INVESTIGACIONES EN INGENIERIA GENÉTICA Y
BIOLOGÍA MOLECULAR (INGEBI)**

27961

A DANIELA

**Por todo eso,...y por eso otro también.
Vos sabés y es nuestro secreto.**

No es por veredas planas que se
sube a las alturas

Séneca.

Dos padres hablaban acerca del porvenir de sus hijos.

Preguntaba uno:

- Qué va a ser de tu hijo cuando termine sus estudios?
- Pues a lo que parece, va a ser un anciano.

Tribune de Chicago.

AGRADECIMIENTOS

Todos tenemos un año en la vida que jamás se olvida, este será el mío sin dudas, por muchas cosas. Basta con decir que escribo estos agradecimientos de la tesis momentos antes de casarme. Para que todo esto junto esté saliendo bien, recibí el apoyo y la ayuda de muchas personas. Nombrar gente en este momento tan especial sería injusto con todos aquellos que me pueda olvidar, por eso no lo voy a hacer, cada uno íntimamente sabe lo que hizo por mí y en que sentido me ayudó, el que no lo hizo también.

Fueron cuatro largos años de trabajo para llegar a este momento y largas horas de insomnio durante las últimas semanas, gracias a todas esas personas, de corazón.

Gracias al Dr. Mariano Levin por el apoyo durante los 4 años en uno de los mejores laboratorios del país. Gracias a la Dra. Mirtha Flawia y al Dr. Héctor Torres por el apoyo durante los 4 años en el INGEBI, uno de los mejores institutos de investigación del país.

Esta tesis fue realizada gracias al apoyo de los siguientes organismos: INSERM-CONICET cooperation program; Projet genome *Trypanosoma cruzi*, INGEBI-CEPH, Ministere d' Affaires Etrangeres (France); Programa Genoma de *Trypanosoma cruzi* asociado al Proyecto Genoma Humano de la Universidad de Buenos Aires; WHO; The United Nations Development Program for Research and Training in Tropical Diseases, the UNIDO-ICGEB-ARG91-01, Fundación Antorchas; Proyecto Genoma de *Trypanosoma cruzi* (CYTED); CONICET, a través de su programa de becas de iniciación y perfeccionamiento.

PUBLICACIONES

El presente trabajo de tesis originó y colaboró en las siguientes publicaciones:

1. Nucleotide cDNA and deduced complete aminoacids sequence of a *Trypanosoma cruzi* P ribosomal protein(P-JL5).
Schijman,A.; Dusetti,N.; Vazquez,M.; Lafon,S.; Levy-Yeyati,P.; Levin,M. IN-
GEBI.
Nucleic Acids Research, Vol 18, 17, 3399, 1990.
2. Nucleotide sequence of a cDNA encoding a *Trypanosoma cruzi* acidic riboso-
mal P1 type protein.
Vazquez,M.; Schijman,A.; Levin,M. INGEBI.
Nucleic Acids Research, Vol 20, 10, 2599, 1992.
3. Nucleotide sequence of cDNA encoding another *Trypanosoma cruzi* acidic
ribosoma P2 type protein (TcP2b).
Vazquez,M.; Schijman,A.; Panebra,A.; Levin,M. INGEBI.
Nucleic Acids Research, VOL 20, 11, 2893, 1992.
4. The *Trypanosoma cruzi* ribosomal P proteins: classification and antigenicity.
Levin, M.; Vazquez, M.; Kaplan, D.; Schijman, A.
Parasitology Today. Vol 9, 10(100), 381-384, 1993
5. The chronic presence of the parasite, and anti-P autoimmunity in Chagas
disease: the *Trypanosoma cruzi* ribosomal P proteins, and their recognition by
the host immune system.
Kaplan,D.; Vazquez,M.; Lafon,S.; Schijman,A.; Levitus,G.; Levin,M.
Biological Research. Vol 26, No 1, 273-277, 1993.
6. Humoral autoimmune response in Chagas disease: *Trypanosoma cruzi* riboso-
mal antigens as immunizing agents
Levin, M.; Kaplan, D.; Ferrari, I.; Arteman,P.; Vazquez, M.; Panebra, A.
Fems Microbiolgy Immunology.

7. Differences in the immune response of patients with Chagas disease or systemic lupus erythematosus to human ribosomal P peptides.
Kaplan, D.; Vazquez, M.; Saunal, H.; Muller, S., Levitus, G., Van Regenmortel, M.H.V.; Levin, M.
Mem. Inst. Oswaldo Cruz, Vol 88, suppl, 51-52, 1993.
 8. Anticuerpos contra las proteínas ribosomales P de *Trypanosoma cruzi* en la enfermedad de Chagas: Autoinmunidad y presencia del parásito.
Levin, M., Kaplan, D., Ferrari, I., Arteman, P., Vazquez, M., Panebra, A.
En el libro "Actualizaciones en la enfermedad de Chagas". Pags. 95-100, 1993.
 9. A short interspersed repetitive element provides a new 3' acceptor site for trans-splicing in certain ribosomal P2 β protein genes of *Trypanosoma cruzi*.
Vazquez, M., Schijman, A., Levin, M.
Molecular and Biochemical Parasitology. 64 (1994) 327-336.
 10. Toward the physical map of the *Trypanosoma cruzi* nuclear genome.
Levin, MJ, Ferrari, I, Vazquez, M, Franco da Silveira, J, Cano, I, Degrave, W, Requena, JM, Alonso, C, Zingales, B, Gonzalez, A, Hernandez, R, Ramirez, JL, Aldao, RR, Saumier, M, Billault, A, Le Paslier, D, Cohen, D.
Mem. Inst. Oswaldo Cruz 89, 17-18, 1994.
 11. Molecular karyotype of clone CL-Brener chosen for the *Trypanosoma cruzi* genome project. Cano, MI, Gruber, A, Vazquez, M, Cortes, A, Levin, MJ, Gonzalez, A, Degrave, W, Rondinelli, E, Zingales, B, Ramirez, JL, Alonso, C, Requena, JM, Franco da Silveira, J.
Molecular and Biochemical Parasitology, 71, 273-278. 1995.
 12. Cloning and sequence analysis of the TcP2 β cDNA variants of *Trypanosoma cruzi*.
Schijman, AG, Vazquez, MP, Ben Dov, C, Ghio, S, Lorenzi, H, Levin, MJ
Biochim. Biophys. Acta, 1264, 15-18. 1995
-

-
13. Detection of polymorphism in the *Trypanosoma cruzi* TcP2 β gene family by Single Strand Conformation Analysis (SSCA).

Vazquez, M.; Beljord, C.; Lorenzi, H.; Bienvenu, T.; Levin, M. J.

Gene, 1995, enviado.

Esta Tesis la realicé de la siguiente manera:

El texto se escribió íntegramente en **MS WORD 7.0** (Microsoft). Las figuras fueron digitalizadas usando un Scanner de página completa **Epson ES-600C** y armadas usando los programas de manejo de imagenes **Paint Shop Pro 3.1**(JASC, inc.) y **Picture Publisher 4.0** (Micrografx, inc.). Los dibujos y diagramas se hicieron usando **Harvard Graphics 4.0** (SPC) , **Visio 3.0** (Shapeware corp.) y **Microsoft Drawing 1.01** (Microsoft).

El ensamblado y formato final de todo se integró usando **MS PUBLISHER 3.0** (Microsoft).

La impresión se hizo usando papel especial EPSON 720 dpi, en una impresora EPSON Stylus Color a 360 y 720 dpi.

Todo el trabajo fue realizado bajo la arquitectura del sistema operativo **WINDOWS 95** (Microsoft) que facilitó las tareas de ensamblado a través de su completa integración con la tecnología OLE (Object Linking and Embedding).

INTRODUCCIÓN

Pag.

El sistema de proteínas ribosomales P.....	1
<i>Trypanosoma cruzi</i> : Generalidades.....	4
<i>Trypanosoma</i> : Características genéticas.....	6

OBJETIVOS.....	12
----------------	----

MATERIALES Y METODOS

Preparación de sondas radiactivas.....	13
Rastreo inmunológico de la bibliotecas en λ gt11.....	14
Hibridizaciones radiactivas.....	16
Manipulación de fagos recombinantes.....	17
Preparación de ácidos nucleicos.....	18
Clonado y subclonado en vectores. Transformación.....	24
Secuenciación de ADN.....	29
Técnicas de manejo de ácidos nucleicos.....	32
Transcripción y traducción <i>in vitro</i>	34
Técnicas especiales de PCR.....	37

CAPITULO 1: Caracterización inicial de proteínas ribosomales P de *T.cruzi*

Revisión Histórica.....	43
Obtención de ADNc que codifican para proteínas ribosomales P de <i>T.cruzi</i>	44

CAPITULO 2: Obtención y análisis de secuencias de ARNm completas de las proteínas ribosomales P .

Puesta a punto y obtención de secuencias de ARNm de proteínas P completas usando la técnica de ARN-PCR en <i>T.cruzi</i>	52
Análisis de las secuencias de ARNm de proteínas P de <i>T.cruzi</i> . Ubicación filogenética y nomenclatura.....	56

CAPITULO 3: Análisis de la expresión de proteínas P

Análisis de las proteínas P por transcripción y traducción <i>in vitro</i> de los ARNm clonados.....	62
Análisis de la expresión de ARNm por clonado usando ARN-PCR y por clonado usando una biblioteca de expresión en λ gt11.....	65
Análisis de la expresión de ARNm por northern blot usando distintas condiciones de temperatura en parásitos <i>in vivo</i>	71

CAPITULO 4: Clonado de genes de proteínas ribosomales P

Organización genómica de proteínas P.....	77
Organización de los loci TcP2 β	86
Polimorfismo de los loci TcP2 β . Análisis por SSCA.....	89
Descubrimiento de una secuencia repetida (SIRE) insertada en un locus TcP2 β : características y consecuencias sobre la expresión.....	95

CAPITULO 5: SIRE: características y organización genómica

Estudio y funcionalidad de SIRE usando como modelo el locus TcP2 β	102
--	-----

Organización y distribución genómicas de SIRE en <i>T.cruzi</i> : estructura del elemento y posible función.....	107
Concepto de SAS (Sitios Asociados a SIRE): su importancia en el proyecto genoma de <i>Trypanosoma cruzi</i>	126
Uso de SIRE en la diferenciación de cepas y en el proyecto genoma: Inter-SIRE y y SIRE-Bubble PCR.....	131

DISCUSION

El genoma de <i>T.cruzi</i> y la organización de genes: el ejemplo de TcP2 β y el resto de las proteínas P.....	136
El genoma de <i>T.cruzi</i> y las secuencias repetidas: el ejemplo de SIRE.....	140

BIBLIOGRAFIA.....	150
-------------------	-----

❖ INTRODUCCION ❖

EL SISTEMA DE PROTEINAS RIBOSOMALES P.

La subunidad mayor del ribosoma contiene una estructura peduncular evolutivamente conservada, involucrada en la unión de factores solubles de traducción y asociada a la actividad GTPasa durante el ciclo de síntesis proteica (1).

Este pedúnculo esta compuesto por una familia de proteínas con características particulares:

- Son las únicas que estan presentes en varias copias por ribosoma.
- Son las únicas proteínas ácidas del ribosoma.

En *E.coli*, el pedúnculo esta formado por un complejo de proteínas L12:L10 en relación 4:1 respectivamente. La proteína L12 es ácida, rica en alaninas (20%) y de bajo peso molecular (12 Kd), presenta una subforma, L7, que solo difiere de L12 en un grupo acetilo en el N- terminal. L10 es también ácida pero de mayor peso molecular (18 Kd) (2). Dos dímeros de L7/L12 están unidos al ARNr 23S a través de la proteína L10 (figura 1)

En las células eucariontes, el pedúnculo se compone de un sistema proteico análogo en propiedades físicas y funciones al de las eubacterias, pero con

características (3).

La proteína L10 es análoga a la proteína eucariota P0 (PM, 38 kd). La disposición estructural en el ribosoma también es conservado: dos dímeros P1/P2 interaccionan con P0 para originar un complejo pentamérico (P1)₂, (P2)₂, P0. No se han encontrado heterodímeros (P1/P2)₂ (4).

A diferencia de sus contrapartes procariotas, las proteínas P son fosforilables. Existe un pool de proteínas P no fosforiladas en el citoplasma, que al ser fosforiladas son capaces de pasar a formar parte del ribosoma y viceversa, las proteínas P unidas a ribosoma al defosforilarse pasan al pool citoplasmático (5).

La secuencia de las proteínas P de células eucariontes presenta un extremo grado de conservación entre las distintas especies (alrededor del 60%). Esta conservación es notable en el extremo C-terminal, donde la secuencia D(D/E)DMGFGLFD-COOH se encuentra en prácticamente todas las proteínas P clonadas, tanto en las de bajo peso molecular (P1 y P2) como en la de mayor peso molecular (P0) (6-13).

La conservación es menor entre las secuencias eucariotas y eubacterianas. Esto se debe a arreglos y translocaciones producidas en el curso de la evolución de las bacterias, puesto que las secuencias del reino de Archaeobacterias se ajustan, aunque con menor homología, a

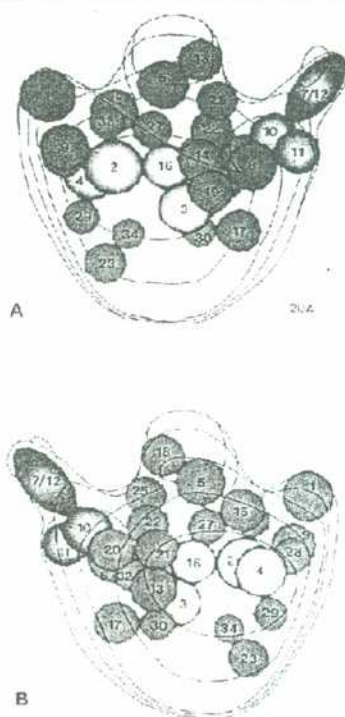


Figura 1- Dos vistas de un modelo gráfico generado por computadora de la subunidad 50 S del ribosoma en *E.coli*.

características diferentes. Por ser ácidas y fosforilables se las designó proteínas A o proteínas P indistintamente (2).

Las células eucariontes tienen dos proteínas P de bajo peso molecular, designadas P1 y P2. Comparten con la proteína procariota pesos moleculares similares, punto isoelectrico, solubilidad en 50% de etanol, y otras

las proteínas P de las células eucariontes (14).

PROTEINAS P1 Y P2

Estas proteínas, análogas a L12, presentan tres regiones bien definidas:

- 1) un dominio globular N-terminal
- 2) una región llamada bisagra rica en alaninas (hasta un 20% de la proteína), con residuos de prolina intercalados.
- 3) Una región C-terminal muy conservada en la evolución, ácida (residuos glutámico y aspártico en su mayoría).

Durante la evolución, en la eubacterias ha habido una translocación del extremo globular N-terminal a la zona C-terminal, dejando a la zona bisagra que en eucariotas se encuentra en la C-terminal en su parte N-terminal (Fig.2) (15).

En eucariontes, L12 esta representada por dos proteínas designadas genericamente tipos I y II. Las proteínas de tipo I tienen un extremo N-terminal un poco mas corto que las de tipo II. La similitud entre ambos tipos es del 25 % aproximadamente(16).

La proteína L12 de Archaeobacterias es codificada por un solo gen como en eubacterias pero se alinea fácilmente con las proteínas eucariontes. Esta proteína archaeobacteriana se parece a las de tipo I (14).

En algunos eucariontes inferiores como la levadura, existen dos genes de tipo I y dos genes de tipo II (16). En los eucariontes superiores y en artrópodos se conoce un solo gen funcional de tipo I (proteína designada

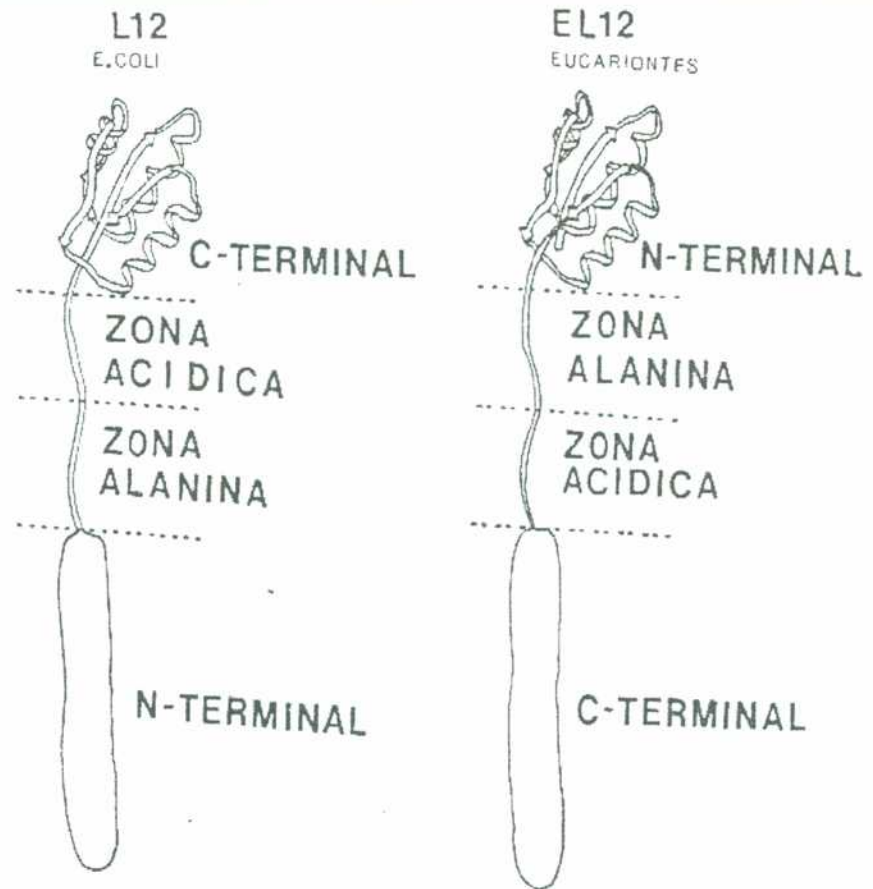


FIGURA 2- Esquema de las proteínas L12 (P2) en *E.coli* y en Eucariontes. Tomado de Liljas y Steitz.

nada P2) y uno de tipo II (proteína designada P1).

En base a la homología que presentan los genes duplicados (55%), la duplicación del gen tipo I y del gen tipo II parece ser reciente. En cambio, la duplicación del gen ancestral que origina los genes tipo I y tipo II es muy antigua, quizás anterior a la separación de los reinos. Luego los linajes de las eubacterias y las archaeobacterias perdieron el gen de tipo II y solo retuvieron el tipo I (17). Este gen sería entonces el que sufrió rearrreglos durante la evolución eubacteriana y dichos eventos ocurrieron probablemente después

de la divergencia entre eubacterias y archaeobacterias eucariontes. El extremo C-terminal conservado en el linaje archaeobacterias-eucariontes no se encuentra conservado eubacterias y su función aún no se ha determinado con exactitud.

PROTEINA P0 (L10)

Una sola copia de esta proteína forma parte del complejo 1:4 que se ensambla en el pedúnculo de la subunidad mayor del ribosoma, tanto en eubacterias como en archaeobacterias y eucariontes. En todos los organismos esta codificada por un solo gen

que da un producto de aproximadamente 35 Kd; presenta la región C-terminal característica de estas proteínas precedida por la región bisagra rica en alaninas (6-13).

En eubacterias, L10 es mas corta debido a un "gap" interno y un truncamiento C-terminal (ver figura 3).

EVOLUCION GENICA

La información funcional, estructural y de secuencia indican que los genes que codifican las actuales proteínas ribosomales P derivan todas de un gen ancestral unico presente en un ancestro primigenio. En 1987, Rich y Steitz, observaron por primera vez que las secuencias de la proteína P0 y las proteínas P1 y P2 humanas exhibían un alto nivel de homología en sus regiones C-terminales: en particular la porción que abarca un modulo de la región globular,

la bisagra, la región acídica y el extremo C-terminal. En algunas especies de archaebacterias, la similitud de secuencia entre L10 y L12 se extiende unos 70 aminoácidos desde el extremo C-terminal hacia el centro de la proteína. Esto sugirió que los genes que codifican estas proteínas de eucariontes y archaebacterias se constituyeron en parte a una secuencia primigenia común (17). De estas comparaciones estructurales se ha interpretado que la proteína L10/P0 pudo haberse originado a partir de la fusión entre la tercera parte de un gen adicional de L12 y la region N-terminal de otro gen, con la consiguiente duplicación de la secuencia modular proveniente de la porción de L12, de dicha fusion génica (17) (Figura 3).

En eubacterias, la similitud de secuencias entre L10 y L12 no es detectable debido al rearrreglo estructural de L12 y a que los dominios bisagra y C-terminales

de L10 están truncados (ver figura 3).

FOSFORILACION DE PROTEINAS P

El blanco de la fosforilación en las proteínas ribosomales P es una serina ubicada en un entorno ácido de reconocimiento para casein kinasas. En *A. salina* se identificó la serina fosforilable (aa98) ubicada en el dominio C-terminal rico en aminoácidos ácidos: EESEEDDD. Esta serina se encuentra conservada en las proteínas P de vertebrados, que presentan en esa región una serina adicional EESEEDDD. En levaduras, el dominio C-terminal ácido K/AEEEDD/ED posee la serina presente en vertebrados y ausente en *Artemia*; sin embargo la fosforilación no ocurre en este residuo, sino en una serina de la porción globular N-terminal (aa19): AAPSAAD (18).

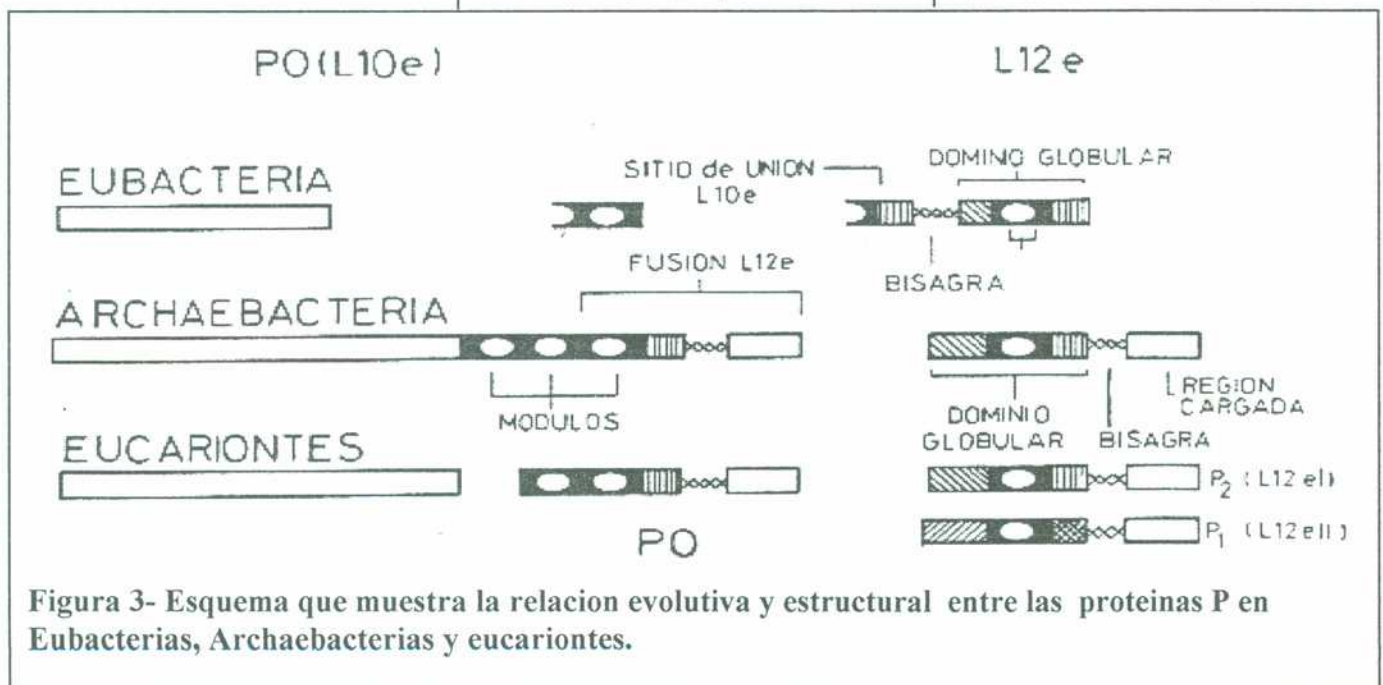


Figura 3- Esquema que muestra la relación evolutiva y estructural entre las proteínas P en Eubacterias, Archaebacterias y eucariontes.

Además de los complejos de proteínas P en ribosomas se han detectado complejos P citosólicos (140 Kd). Estos complejos citoplasmáticos presentan menor grado de fosforilación que los complejos incorporados a ribosomas. La fosforilación parece aumentar la actividad de las proteínas ácidas (19) y su

afinidad por los ribosomas. A partir de estos hallazgos se ha propuesto que la fosforilación podría ser indispensable para la integración de estas proteínas al ribosoma.

Si bien la fosforilación ha sido descrita en varios organismos como *A. salina*, *S. cerevisiae*, *R. norvegicus*, y *H. sapiens*, no sería

un fenómeno universal dentro de los eucariontes, ya que se ha demostrado que las proteínas homólogas del protozoo *Tetrahymena pyriformis* no se fosforilan in vivo (20). En *Tetrahymena* la región C-terminal ácida no presenta serinas (20).

TRYPANOSOMA CRUZI : GENERALIDADES.

CARACTERISTICAS DEL GENERO

El género *Trypanosoma* comprende parásitos digenéticos, es decir que alternan su ciclo de vida entre dos hospedadores.

Incluye dos grupos de organismos:
-**Grupo Salivaria:** constituido por un conjunto de especies

patógenas, que se multiplican en el hospedador mamífero como tripomastigotes, completan su desarrollo en la glándula salival del vector y son transmitidos por inoculación con las secreciones salivares. (ej. *T. brucei*).

-**Grupo Estercoraria:** constituido por un conjunto de especies no patógenas (con excepción de *T. cruzi*), que se multiplican en el

mamífero como amastigotes, completan su desarrollo en el intestino del vector y son transmitidos por contaminación con las heces.

CARACTERISTICAS DE LA ESPECIE.

Trypanosoma cruzi presenta 3 estadios, distinguibles según las características del flagelo y la posición relativa del kinetoplasto respecto al núcleo celular.

- **Epimastigote:** es la forma flagelada replicativa en el intestino medio del insecto vector. El flagelo emerge de la región lateral anterior respecto a la dirección del movimiento del organismo y el kinetoplasto es anterior al núcleo celular.

- **Tripomastigote:** es la forma flagelada no replicativa. El flagelo emerge del extremo apical posterior y el kinetoplasto es posterior al núcleo celular. Se llama tripomastigote sanguíneo en el mamífero y metacíclico a la forma diferenciada infectiva en el insecto. Los tripomastigotes sanguíneos pueden infectar macrófagos por un evento de

REINO	PROTISTA
PHYLUM	SARCOMASTIGOPHORA
SUBPHYLUM	MASTIGOPHORA
CLASE	ZOOMASTIGOPHORA
ORDEN	KINETOPLASTIDA
SUBORDEN	TRYPANOSOMATINA
FAMILIA	TRYPANOSOMATIDAE
GENERO	TRYPANOSOMA
SUBGENERO	SCHIZOTRYPANUM
ESPECIE	CRUZI

penetración activa y otros tipos celulares por un evento dirigido por el parásito (21).

Amastigotes: es la forma no flagelada, de localización intracelular en vacuolas ácidas y replicativa en el citoplasma del hospedador.

En base a distintos estudios realizados como por ejemplo: Zimodemas (caracterización de perfiles isoenzimáticos), Ezquizodemas (fingerprinting de minicírculos), patogenicidad en ratones, etc, se ha descripto que poblaciones naturales de *T. cruzi* presentan una composición heterogénea (22). Algunas de las cepas de laboratorio mas comunes son Y, RA, Tulahuen, Maracay y la recientemente denominada CL Brener que es la cepa elegida para el proyecto genoma de *T. cruzi*

ENFERMEDAD DE CHAGAS

La enfermedad de Chagas se desarrolla en dos fases:

-Fase aguda: El período agudo es en general asintomático, se estima que mas del 90 % de lo individuos infectados pasa a la fase crónica sin advertirlo (23).

El conjunto de signos clínicos que caracterizan la fase aguda comprende: fiebre, parasitemia detectable, linfadenopatía, hepatoesplenomegalia, vómitos, diarrea y especialmente los signos de puerta de entrada, como el ojo de Romaña. En un bajo numero de casos (1-5 %) se observa la aparición de una miocarditis, generalmente

reversible (23).

-Fase crónica: La mayor parte de la población crónicamente infectada no presenta evidencias discernibles de enfermedad; solo un porcentaje de hasta un 30% dependiendo de la región desarrolla al cabo de 10, 20 o más años, síntomas y/o signos de daño visceral. Esto es lo que se denomina enfermedad de Chagas crónica, cuya manifestación clínica mas frecuente en Argentina es una cardiopatía (24).

En la cardiopatía chagásica crónica, desde los estadios iniciales hasta los cuadros terminales, se detectan focos inflamatorios activos, los que podrían desempeñar un rol preponderante en la destrucción lenta y progresiva del miocardio.

Sin embargo, el escaso número de parásitos y el tipo de células halladas en los infiltrados de las lesiones crónicas han llevado a proponer que la patología chagásica crónica se desencadena como consecuencia de un proceso autoinmune (24).

Este fenómeno autoinmune sigue siendo hoy en día tema de debate. Por un lado se ha propuesto que los parásitos causan un desequilibrio en la regulación del sistema inmune, promoviendo la expansión policlonal de linfocitos B con consiguiente producción de autoanticuerpos naturales.

Por otro lado, se explica el origen la autoinmunidad a través de un proceso del mimetismo molecular

PROTEINAS P Y ENFERMEDAD DE CHAGAS CRONICA.

Las proteínas ribosomales P humanas son antígenos en Lupus Eritematoso Sistémico (LES) (26). El 12 % de pacientes con LES, que incluye una subpoblación con manifestaciones de psicosis, compromiso muscular y cardíaco, presenta autoanticuerpos anti-P.

La reactividad anti-P de estos sueros estaba dirigida hacia un epítopo constituido por los 11 residuos C-terminales de las proteínas P (26, 27).

El suero de pacientes chagásicos crónicos presenta un elevado título de anticuerpos contra proteínas parasitarias intracelulares. Se han descripto anticuerpos contra proteínas flagelares, proteínas de citoesqueleto y citoplasmáticas (28). En particular, los enfermos chagásicos crónicos con cardiopatía se caracterizan por tener un alto nivel de anticuerpos contra las proteínas ribosomales P del parásito y esta reactividad anti P podría estar relacionada con las ya mencionadas características autoinmune de las lesiones cardíacas chagásicas.

En nuestro laboratorio, se rastreo una biblioteca de ADNc de expresión del parásito con estos sueros y se aisló el recombinante JL5, entre otros.

La identificación de la secuencia de JL5 demostró que se trataba de una proteína ribosomal P del parásito y confirmó la conservación del autoepítopo lúpico (29, 30, 31)

La naturaleza autoreactiva de los anticuerpos dirigidos hacia la región C-terminal de la proteína

P de *T.cruzi*, fue demostrada en experiencias de Western blot con extractos de ribosomas humanos y con proteínas de fusión beta-galactosidasa-proteínas P1 y P2 humanas.

Esto llevo a realizar un análisis detallado de las proteínas P del parásito, en electroforesis bidimensionales y posterior Western blot, indicando que el sistema P de *T.cruzi* es uno de los mas complejos descritos hasta ahora (31).

En la figura 4, se observan alrededor de 10 proteínas P de bajo peso molecular, identificadas como tales por su reacción con los anticuerpos anti-P lúpicos que reaccionan con sus secuencias C-terminales: DDDMGFGLFD. Es sorprendente que en otras infecciones crónicas causadas por protozoarios cercanos evolutivamente a *T.cruzi* la especificidad anti-P no sea detectada.

La construcción de un péptido con los ultimos 13 aminoácidos C-terminales de la proteína ribosomal JL5, llamado R13, permitió delimitar la especificidad compartida por los anticuerpos anti-JL5 de sueros chagásicos y por los anti-P de sueros lúpicos (31). El perfil de isotipos anti-R13, de sueros chagásicos crónicos, coincidió con el perfil de anticuerpos anti-*T.cruzi* y fue distinto a la distribución isotípica anti-R13 de los sueros lúpicos (31). Estas experiencias sugirieron que en la cardiopatía chagásica crónica, los autoanticuerpos anti-R13 se originarían por un mecanismo de mimetismo molecular.

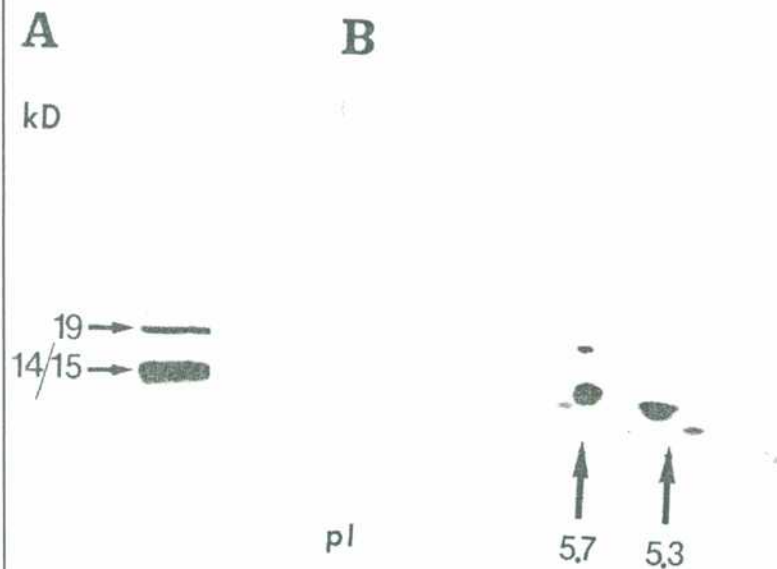


Figura 4- Parte A, Western blot de ribosomas de *T.cruzi*. Parte B, gel bidimensional de los mismos ribosomas revelado con Ac. anti P

TRYPANOSOMA : CARACTERISTICAS GENETICAS

El modelo mas estudiado genéticamente es el tripanosoma africano (*T.brucei*). Por esa causa, muchos de los mecanismos y conceptos explicados en esta sección fueron originalmente estudiados en *T.brucei* aunque la extrapolación a los tripanosomátidos en general se acepta como válida.

Los kinetoplástidos se encuentran entre las células eucariotas mas divergentes que se conocen. Quizás como consecuencia de su extensa separación de la linea principal de los eucariotas poseen algunas propiedades biológicas que no comparten con la mayoría de los organismos. Entre estas se pueden citar (32):

- Variación antigénica de los tripanosomas africanos.
- Ausencia de intrones
- Trans-splicing en lugar cis-splicing.

- Transcripción policistronica.
- K-DNA mitocondrial (una de las estructuras de ADN mas inusuales que se conocen).
- Edición de ARN (RNA Editing).
- Mini ARN ribosomales en mitocondria.
- Alta plasticidad de cariotipo.
- Antigenos con repeticiones.
- Glicosomas (único organismo que compartimentaliza las enzimas glicolíticas)
- Nucleótidos modificados en el ADN genómico.

Se conocen dos formas de relación parásito-hospedador:

- 1) Monogénicos, que cumplen su ciclo de vida dentro de un solo hospedador invertebrado.
- 2) Digenéticos, que cumplen su ciclo de vida alternando entre dos hospedadores, generalmente uno invertebrado y otro vertebrado.

Las 123 especies monogénicas

conocidas hasta ahora pertenecen a los géneros *Crithidia*, *Herpetomonas*, *Blastocrithidia* y *Leptomonas*. Son parásitos en artrópodos y se los conoce como tripanosomátidos inferiores porque los géneros digenéticos como *Leishmania* y *Trypanosoma* han evolucionado probablemente de un ancestro monogenético (33).

TRANS-SPLICING

Los genes eucariotas se organizan de un modo diferente que en los organismos procariotas.

Se suponía que el concepto de operón, organización y expresión de varios genes relacionados bajo la dirección de secuencias regulatorias comunes, no se aplica en eucariotes y los ARNm policistrónicos no se observan en estos organismos.

Este no es el caso del orden de los kinetoplástidos.

Muchos genes se encuentran organizados en tandem (GADPH, Calmodulina, Tubulina, Ubiquitina, HSP70, Actina, PARP) (33-36), separados por regiones intergénicas muy cortas que, se supone, carecen de señales para la iniciación de la transcripción. Más aún, muchos genes, aparentemente, no relacionados entre si se organizan de esta forma compacta y también carecerían de estas señales. En realidad, se sabe que estos organismos forman unidades transcripcionales muy largas utilizando un único promotor ubicado al principio de dicha unidad transcribiendo ARNs

precursores policistrónicos de gran longitud (37). Por otra parte, todos los genes clonados y estudiados carecen de intrones, esto hace suponer que todo el genoma nuclear de estos parásitos carece de intrones.

Dado que los ribosomas eucariotes no son capaces de reconocer sitios internos AUG de inicio de traducción en estos ARN policistrónicos, el ARN mensajero maduro se forma por un mecanismo absolutamente novedoso: **Trans-splicing**. (38)

Durante el estudio de la variación antigénica en *T.brucei*, se encontro que todos los mensajeros de VSGs (Variant Surface Glicoprotein) tenían una secuencia líder común en el extremo 5' que no se encontraba codificada en el gen. Posteriormente se observó que esta secuencia estaba presente en todos los ARNm clonados de distintas proteínas. Esta secuencia se describió en un principio con una longitud de 35 nucleótidos (ahora se sabe que tiene 39 nucleótidos, ver mas adelante). Boothroyd y Cross le dieron el nombre de "Miniexón" en 1986. Este miniexón, también conocido como SL (Spliced Leader), esta codificado por un exón separado que se une a los mensajeros por una reacción en trans. Hay aproximadamente 200 copias de SL ordenadas en "clusters" en el genoma, cada SL forma parte de una secuencia de 1.35 Kb que esta repetida en tandem. La unidad de 1.35 Kb tiene poca conservación entre los distintos kinetoplástidos, con la excepción del SL en si mismo que se encuentra

altamente conservado. El único transcripto detectable de esta unidad es un ARN de 140 nucleótidos que lleva al SL en el extremo 5' (38). Este transcripto se conoce como medRNA (miniexon derived RNA) y aun se discute si es transcripto por ARN polimerasa II o III. Su tasa de recambio es muy alta (alrededor de 4 minutos IN VIVO) (39).

Donelson, en 1986, demostro que todos los ARNm maduros de kinetoplástidos llevan esta secuencia líder. Basándose en un ensayo de traducción IN VITRO conocido como "hybrid arrested translation", en el cual utilizó un oligonucleótido antisentido del SL y ARN mensajero purificado de distintos kinetoplástidos, observó que no se traducía ninguna proteína (40). Mientras que en el control en cual no se agregaba el SL antisentido, la traducción era normal.

Por el término Trans-splicing, entonces se entiende la maduración de ARN precursores policistrónicos a través de la unión de transcriptos independientes en una reacción bimolecular que promueve la adquisición por parte del mensajero maduro de una secuencia líder de 39 bases.

El transcripto del SL puede ser dividido en dos dominios funcionales separados por el consenso del sitio dador de splicing: el dominio 5' que contiene al SL o miniexón y el dominio 3' que es variable en longitud.

Durante la reacción de Trans-splicing, se produce el clivaje en el sitio dador 5' del SL, seguido de

una trans-esterificación del fosfato 5' del minitrón de 100 bases en una unión fosfodiéster 2'-5' creando una configuración del ARN en forma de Y. La subsecuente trans-esterificación con el grupo OH 3' del SL, en el sitio de splicing 3', libera la estructura en Y completando la reacción. (ver figura 5) (38). Esta estructura en Y se conoce como sitio de bifurcación (branchpoint).

El armado del Spliceosoma difiere del conocido para la reacción de Cis-splicing.

En primer término la ribonucleoproteína U1, que se une al dador 5' en cis-splicing, no se ha encontrado en tripanosomas. Tampoco se ha encontrado U5, que reconoce el sitio 3' de cis-splicing. Si se han identificado U2, U4 y U6, con la particularidad que a U2 le falta el sitio de reconocimiento del branch site y el stem loop III (38).

Sin embargo, Bruzik y Steitz demostraron en 1990 que el SL podía sustituir a la partícula U1 en un sistema IN VITRO de cis-splicing mamífero, quizás debido a una estructura secundaria similar, a pesar de la disparidad en las secuencias (41).

La ausencia de U5 quizás se justifique por la observación de que en Trans-splicing no hay un único sitio aceptor 3' sino varios dentro de cierta distancia la iniciación de la traducción.

Otro fenómeno asociado es el de trans-capping, debido a que es el SL el que lleva la estructura CAP 5' y la transfiere junto con el en durante la reacción de Trans-splicing.

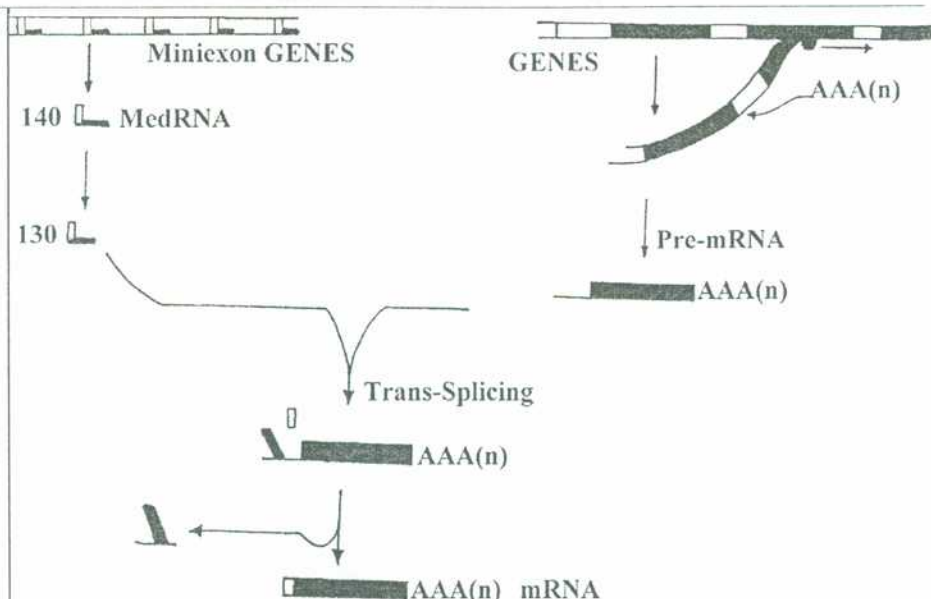


Figura 5- Esquema de la reacción de Trans-splicing en Tripanosomas

Esta estructura de CAP es muy inusual en eucariontes e involucra las 4 primeras bases del SL; razón por la cual en principio se lo describió con una longitud de 35 bases, dado que las 4 primeras se encuentran altamente modificadas.

La estructura de CAP es la siguiente:

**m7G(5')ppp(5')m6/2AmpAmp
Cmpm3UmpAAGC...**

En 1990, Nina Agabian demuestra que la inhibición de la metilación anula el mecanismo de Trans-splicing, al imposibilitar la formación de la estructura CAP (42).

El mecanismo de Trans-splicing cumple así un papel fundamental en el procesamiento de ARN policistrónicos en ARNm maduros. La extrema conservación de la secuencia del SL se podría explicar teniendo en cuenta que cumpliría la función

de U1; sin embargo, nada de esto explica como se obtienen niveles tan diferentes a veces de ARNm maduros a partir de genes que se transcriben dentro de una misma unidad transcripcional a partir de un mismo promotor.

Esto llevo a proponer que la reacción de Trans-splicing involucra también algún mecanismo de control sobre los niveles finales de ARNm, es decir un mecanismo post-transcripcional.

Dos grupos trabajaron independientemente sobre esta idea, el grupo de Dyann Wirth y el grupo de Stephen Beverly (43, 44). El grupo de Wirth demuestra en 1992 que dicho mecanismo de regulación post-transcripcional ejercido por el trans-splicing existe. Este grupo estudia las regiones intergénicas de los genes tubulina. Una zona intergénica en kinetoplástidos involucra 2 tipos de señales: la señal de poliadenilación (cuyo consenso no se conoce) para el gen anterior

y la señal de Trans-Splicing para el gen posterior. La señal de Trans-Splicing involucra una región de polipirimidinas de longitud variable, dentro de la cual en general se mantiene básicamente la secuencia:

TTTCpy (4-40 bases) AG.

Este grupo demuestra que al insertar en la zona intergénica más regiones de polipirimidinas repetidas se aumenta proporcionalmente la cantidad de ARNm de tubulina en el citoplasma (43).

Estos resultados indican que estas regiones contienen señales que afectan la estabilidad del mensajero, y que son secuencias involucradas en el Trans-Splicing.

El grupo de Beverly publica en 1993 un trabajo en el cual demuestra que el Trans-Splicing se encuentra acoplado a la poliadenilación, basándose en la proximidad de ambas señales dentro de la zona intergénica y en la falta de una señal consenso de poliadenilación.

Ellos observan que el mecanismo y la regulación de las reacciones de procesamiento del ARNm que ocurren en las regiones intergénicas de los policistrones es el determinante primario de la regulación de la expresión. Esto contrasta con la mayoría de los organismos eucariotas en donde la iniciación de la transcripción es el principal determinante (44).

Así el sitio de poliadenilación es determinado por la señal de Trans-Splicing del gen posterior, evento que ocurre luego del armado del

Ensamblado de Spliceosoma

Unión del complejo Poli A

Scanning 3'-5'

Clivaje

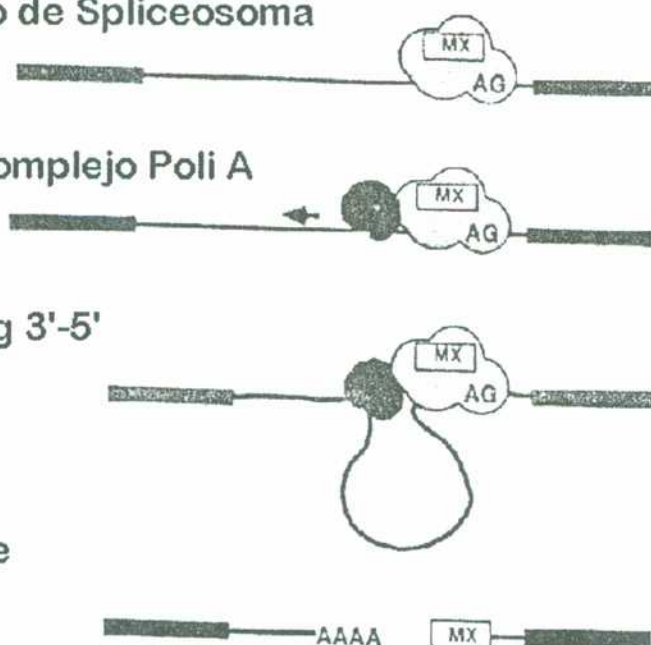


FIGURA 6- Acople de la reacción de Trans-Splicing y Poliadenilación

spliceosoma y a través de un scanning 3'-5'. El corte para poliadenilación se produce dentro de ciertas "ventanas" permitidas, marcadas por regiones ricas en (T)nA. De esta forma, una sola señal en la región intergénica controla el Trans-Splicing y la Poliadenilación y regula los niveles finales de ARNm (mecanismo post-transcripcional) (44).

El mecanismo de Trans-Splicing, sin embargo, no es exclusivo de los Kinetoplastidos. Se ha descrito también en nematodos y trematodos (ej. *C.elegans*) y en el alga *Euglena*. El mecanismo en estos organismos presenta varias diferencias con el de tripanosomas y en la actualidad se discute si son mecanismos o si coevolucionaron en forma independiente. Algunas diferencias son las siguientes

(45):

- Coexisten Cis y Trans-splicing
- Solo un 15 - 20% de los ARNm utilizan el Trans-splicing.
- Tienen 2 tipos de SL: SL1 y SL2
- No tienen homología con el SL de tripanosomas.
- Sobre un mismo transcripto pueden hacer Cis y Trans-splicing
- Poseen todas las UsnRNP (U1, U2, U4, U5, U6)
- SL2 es usado solo cuando hay transcripción policistronica.

VARIACION ANTIGENICA Y REGULACION DE EXPRESION GENETICA.

Los tripanosomas, en general, no se rigen por las mismas reglas que controlan la expresión génica en el resto de los eucariontes. Un ejemplo de ello es el mecanismo de Trans-Splicing explicado mas arriba. Otros provienen del estudio del mecanismo de la variación antigénica en el tripanosoma africano (*T.brucei*).

La habilidad que tienen los tripanosomas africanos para sobrevivir y replicarse en el torrente sanguíneo de los mamíferos se debe a su cubierta de Glicoproteínas Variables de Superficie (VSG), las cuales, como su nombre lo indica, son capaces de variar antigénicamente, permitiéndole al parásito evadir la respuesta inmune del hospedador. La frecuencia con la cual aparece una nueva cubierta de VSGs en superficie va desde 10^{-2} a 10^{-6} por división. Como resultado, las infecciones muestran picos sucesivos de parasitemia, cada uno correspondiente a un tripanosoma con distinto antígeno VSG no enfrentado previamente por el sistema inmune (46).

El tripanosoma procíclico que vive en el insecto vector no expresa genes de VSG; en lugar de ello, produce cantidades abundantes de una proteína invariable que recubre la superficie. Esta proteína a recibido dos nombres, Prociclina o **PARP** (por Proteína Acida

Repetitiva de Procíclicos). Cuando el procíclico se convierte a la forma infectiva, se deja de expresar rápidamente PARP y se expresa un repertorio limitado de 14 tipos diferentes de VSGs, entre más de 1000 posibles, para preparar la entrada al hospedador mamífero.

El mecanismo por el cual se logra el "switch" de la expresión de distintos genes VSG es muy complejo, a la vez que no del todo bien entendido, y no se profundizará más en esta introducción (46).

La expresión de los genes VSG es telomérica y forma unidades transcripcionales policistrónicas muy largas (de un promedio de 50 a 60 Kb). Esta unidad transcripcional codifica hasta 7 genes asociados a las VSG (conocidos por sus siglas en inglés como **ESAGs**). Los ESAGs son genes que codifican proteínas invariables, algunos de los cuales pudieron ser identificados, como por ejemplo: una proteína que une transferrina e incluso una adenilato ciclase (46, 47, 48).

Se ha investigado durante mucho tiempo como el tripanosoma logra controlar esta expresión estadio específica de VSGs y PARPs. Los resultados se vieron demorados por la dificultad para identificar promotores correspondientes a estos genes, dada la naturaleza policistrónica y la gran longitud de estas unidades transcripcionales.

Sorpresivamente se encontró que la transcripción de dichas unidades era resistente al tratamiento con α -amanitina, droga que como se sabe inhibe a la

ARN polimerasa II, enzima encargada de la transcripción de los genes que codifican para proteínas. Esta resistencia es típica de ARN polimerasa I, enzima que transcribe solo los ARN ribosomales. Este caso es único dentro de los eucariontes.

Se han planteado para este hecho distintas hipótesis:

- Que son transcritos por una forma modificada de ARN polimerasa II
- Que son transcritos por la ARN polimerasa I
- Que son transcritos por una nueva ARN polimerasa (ARN polimerasa IV)

Cualquiera sea la opción correcta, este hecho ya marca una clara diferenciación entre la expresión de estos genes VSG y PARP y el resto los genes que codifican para proteínas.

Distintas evidencias se han mostrado en favor de la transcripción por ARN polimerasa II modificada (49):

- Los tripanosomas que realizan variación antigénica son los únicos organismos conocidos que tienen dos genes para ARN polimerasa II (incluso *T.cruzi* tiene un solo gen). La homología entre estos dos genes es muy alta.
- La sensibilidad a 1,10 phenantroline y Mn^{++} es típica de una ARN polimerasa II.

Las evidencias en favor de una ARN polimerasa I son las siguientes:

- α -amanitina resistencia
- La resistencia a sarkosyl también típica de polimerasa I (50).
- El mapeo de un promotor en la unidad de VSG118 que comparte características con promotores de

ARN polimerasa I de eucariontes.

- Construcciones hechas con el promotor de ARN ribosomal de tripanosoma logran dirigir la expresión de genes que codifican para proteínas en este organismo (51).

- Se realizó el Knock-out de los dos alelos correspondientes al segundo gen de polimerasa II sin ningún efecto sobre la viabilidad o transcripción de genes VSG (52).

- Algunos, pero no todos, genes VSG se observaron transcribiéndose en el nucleólo.

La imposibilidad de la ARN polimerasa I para transcribir genes para proteínas en eucariontes es debida a que no es capaz de agregar un CAP 5' en el ARNm. Esto no sería un problema en tripanosomas debido a que el CAP es agregado en trans con la adición de SL.

Sin embargo, mas allá de cual sea la polimerasa que transcribe estas unidades, se sabe que los promotores de las unidades VSG y PARP son constitutivos (53). La pregunta es cómo logra el tripanosoma el apagado (shut off) y encendido tan rápido de estos genes al cambiar de estadio.

Se ha observado que durante el pasaje del hospedador mamífero (37°C) al insecto vector (28°C), la transcripción de los genes VSG se frena a causa de un descarrilamiento prematuro de la ARN polimerasa del templado por la reducción de la temperatura. Como el gen VSG es el último de la unidad de 60 Kb de longitud, la polimerasa no llega a transcribirlo durante el estadio a 28°C a pesar de que el promotor de la unidad

continúa activo (48).

En los genes PARP, la regulación se relacionaría más con señales presentes en la zona 3' no codificante (54). Estas señales serían las responsables de la rápida degradación de los ARNm de PARP cuando el parásito entra en la sangre del hospedador mamífero. Una secuencia de 16 bases que forma parte de una estructura en "stem-loop" se conserva entre los genes PARP de distintas especies. Un estudio detallado determinó que no tenía efecto sobre el procesamiento del ARNm, su estabilidad, o el transporte, pero sí se requería para una eficiente traducción en las formas procíclicas (54, 55).

Hay varios niveles novedosos en los que se ejerce la regulación de expresión de estos genes estadio específicos en tripanosoma, pero nunca es a nivel de la iniciación de la transcripción (55).

KINETOPLASTO Y RNA EDITING

Los kinetoplástidos deben su nombre al **kinetoplasto** (kDNA). Este DNA forma parte de la única mitocondria que tienen estos organismos protozoarios.

El kDNA consiste de aproximadamente 5000 minicírculos (~1.5 Kb cada uno) y de 25 a 50 maxicírculos (37 Kb cada uno) que están topológicamente concatenados en una gran red. (56)

Cuando el parásito entra en el torrente sanguíneo del hospedador mamífero, la mitocondria se

degenera transformándose en un canal periférico. Los citocromos están ausentes y deja de funcionar el ciclo de Krebs. Por lo tanto, la energía se obtiene solo por fosforilación a nivel de sustrato (glicólisis) que es llevada a cabo muy eficientemente por una organela especializada única en los eucariontes: el **glicosoma**. (57)

Cuando el parásito pasa al insecto vector, la mitocondria recupera su tamaño y se vuelve activa, obteniéndose la energía a través del proceso de respiración. Sin embargo, no se ha podido demostrar que exista un proceso de traducción activo en estas mitocondrias: no se pudieron identificar los tRNAs y ribosomas, aunque sí se encontraron los ARN ribosomales. Estos ARN ribosomales se encuentran entre los más divergentes conocidos y son mucho más chicos en tamaño (9S y 12S) comparados con los de *E.coli* (16S y 23S).

Debido a su estructura en red, el kDNA requiere un novedoso mecanismo de replicación. Esta replicación ocurre al mismo tiempo que la división nuclear, durante una fase S corta. Los minicírculos se encuentran covalentemente cerrados y concatenados en la red (forma I). Durante la fase S son liberados de la red por una topoisomerasa II, y los minicírculos libres se replican en una estructura ϕ . La progenie de minicírculos, que están "nickeados", se van enganchando en la periferia de la red en replicación por la acción de otra topoisomerasa. Como el enganche está restringido a la periferia, la red en replicación desarrolla dos

zonas. La zona central contiene minicírculos covalentemente cerrados y aún no replicados, y la periferia contiene minicírculos "nickeados" que ya sufrieron replicación. Cuando la fase S se termina, la red contiene casi 10000 minicírculos (forma II). Los minicírculos son reparados y concatenados formando dos redes iguales que son segregadas durante la división celular (58, 59). Se desconoce el propósito del concatenamiento en red. El mecanismo en detalle es más complicado y puede consultarse en las citas mencionadas.

Los maxicírculos corresponden al el genoma mitocondrial de eucariontes y los minicírculos tienen una función desconocida. Sin embargo, Benne et al., en 1986, observaron que si bien podían purificar la subunidad III de la citocromo oxidasa (COIII) y detectar ARNm correspondiente a la misma, no podían identificar el gen en el kDNA (tampoco existía una copia nuclear). Finalmente, un análisis detallado de secuencias de maxicírculo, mostró una secuencia lejanamente parecida al gen de COIII, las diferencias con respecto al ARNm clonado se debían exclusivamente a la delección o inserción de uridinas (U) en el ARNm. A este mecanismo se le dio el nombre de **RNA Editing** (edición de ARN). Muchos genes codificados en maxicírculo pasan por este proceso, y algunos llegan a ser editados hasta en un 50%. Distintas evidencias muestran que el mecanismo procede en dirección 3'-5' (56), pero el principal problema que enfrentaron los biólogos moleculares era determinar cual

podía ser el templado que dirigía esta edición, de hecho durante mucho tiempo se pensó que el templado no existía. Nancy Maizels y Alan Weiner escribieron en 1988 en Nature en un artículo titulado "En busca del templado" (60): "El dogma central de la biología molecular ha sobrevivido a tiempos difíciles, pero esta adición de secuencias sitio específica y no templada podría llegar a darle un golpe mortal".

Finalmente, la situación se resolvió con el descubrimiento en 1990 de los "ARN Guías" (gRNAs) (56). Los gRNAs son codificados tanto por los maxicírculos como por los minicírculos y son los encargados de guiar la edición del ARN en mitocondrias. En la actualidad este mecanismo se encuentra ampliamente estudiado y las reacciones de edición se han logrado llevar a cabo IN VITRO.

Todos estos inusuales aspectos de la genética de los tripanosomas los convierten en uno de los mas divergentes y a la vez fascinantes organismos unicelulares dentro del reino de los Eucariotas.

OBJETIVOS DE LA PRESENTE TESIS

Los objetivos de la presente tesis fueron los siguientes:

- Caracterizar la familia de proteínas ribosomales P de *T. cruzi* a nivel de organización genómica y analizar la expresión de los genes que la componen.
- Utilizar esta familia génica, una vez caracterizada, como ejemplo para entender aspectos moleculares de la organización y expresión de genes en el parásito.
- Análisis de la organización e influencia de las secuencias repetidas en el genoma y su importancia en el desarrollo del Proyecto Genoma de *Trypanosoma cruzi*.

MATERIALES Y MÉTODOS

1. PREPARACIÓN DE SONDAS RADIOACTIVAS:

Se utilizó el Método de marcación por cebado al azar (*random Priming*) empleando el kit de *multiprime DNA labelling system* (Amersham) RPN.1601 Y/Z o BPN 1601 Y/Z, siguiendo las indicaciones de los manuales adjuntos.

Las reacciones de marcado fueron realizadas con 25 a 50 ng de ADN en 10 μ l de H₂O, a partir de insertos amplificados por PCR, de insertos electroeluidos de geles de agarosa, o bien de insertos incluidos en taquitos de agarosa de bajo punto de fusión.

Las muestras fueron desnaturalizadas a 95-100°C, por 2-5 min y sumergidas de inmediato en hielo. De acuerdo al deoxinucleótido radiactivo utilizado, dCTP [α P³²] 3000 Ci/mmol, o dATP [α P³²] 3000 Ci/mmol, se emplearon distintos kits en las siguientes condiciones:

Deoxinucleótido radiactivo	dATP α P ³² (5 μ l)	dCTP α P ³² (5 μ l)
Buffer de reacción/dNTP mix	-	5-10 μ l
Buffer de reacción	2,5/5 μ l	-
dCTP – dGTP – dTTP frios	2/4 μ l c/u	-
Random hexamers/BSA	2,5/5 μ l	2,5/5 μ l
H ₂ O csp 25/50 μ l	x μ l	x μ l
dNTP α P ³² 10 μ Ci/ μ l	2,5/5 μ l	2,5-5 μ l
Klenow 1 U/ μ l	2 μ l	2 μ l

Los componentes se mezclaron por spin down y la reacción se dejó no menos de 5 hs a temperatura ambiente. Concluida la reacción de marcado se agregó 5 μ l de ADN carrier, un vol de 4M AcONH₄, pH: 4.5 y 4 vol de Etanol absoluto. Se dejó 15 minutos a -70°C. Se calentó el tubo a 37°C por 2 min, para resuspender los nucleótidos libres que pudieron haber precipitado.

La muestra se centrifugó a 12000 rpm en microcentrifuga durante 15 min. El pellet se lavó en 0,5 ml de una solución de 0,67 N AcONH₄, 67 % de etanol, a temp. amb., por centrifugación. Se secó al vacío, se estimó la marca con contador Geiger. La sonda se resuspendió en 100-200 μ l de

buffer TE y se desnaturalizó a 95-100°C por 5 min, inmediatamente antes de ser usada.

En los casos de sondas preparadas a partir de agarosa de bajo punto de fusión, se tomó el taco de agarosa, con aprox. 250 ng del fragmento a marcar, el cual fue transferido a un tubo Eppendorf. Se agregó 3ml H₂O por gramo de agarosa.

Se hirvió durante 7 min para fundir la agarosa y desnaturalizar el ADN. La muestra fue alicuotada en fracciones de 25 µl, con 25-50 ng de ADN por fracción que se conservó a -20°C. Cada fracción fue descongelada una sola vez; fue calentada a 100°C por 1 min, transferida a un baño de 37°C por 10 min y marcada según el procedimiento recién descrito.

2. RASTREO INNUNOLÓGICO DE LAS BIBLIOTECAS EN λ gt11.

Preparación de bacterias *E. coli* RY1090 para infección con λ gt11.

Se inoculó una colonia aislada de *E. coli* RY 1090 en 100 ml de LB, 50 µg/ml ampicilina, 0,2% maltosa y se incubó durante la noche a 37°C con agitación. Las bacterias se cosecharon por centrifugación a 4000 rpm durante 5 min a 4°C, y se resuspendieron en 0,4 vol de MgSO₄ 10 mM estéril y helado. Las bacterias preparadas se conservaron a 4°C por no más de 15 días.

Plaqueo de la biblioteca de λ gt11 en *E. coli* RY 1090

Aproximadamente $4 \cdot 10^4$ ufp en 100 µl de SM se mezclaron con 600 µl de bacterias competentes (ver paso anterior) y se incubaron durante 20 min a 37°C. A continuación se mezclaron con 3,5 ml de LB-top-agarosa fundido a 48°C, se homogeneizaron y vertieron sobre cajas de Petri conteniendo LB-ampicilina previamente mantenidas a 42°C.

Se incubó a 42°C hasta que las playas de lisis fueron claramente visibles (alrededor de 3 hs).

Transferencia de las proteínas de fusión producidas por los fagos recombinantes a nitrocelulosa

Se colocó sobre las placas líticas un disco de nitrocelulosa (Schleicher and Schuell BA85) o Hybond C (Amersham), previamente embebido en una solución de IPTG 10 mM y secado sobre papel de filtro (se puso en contacto con los fagos la cara del disco que no estuvo en contacto con el papel).

La posición del filtro de nitrocelulosa se marcó con una aguja o alfiler embebido en tinta china, atravesando el filtro y el agar. Se hicieron tres marcas asimétricas de uno, dos y tres puntos.

Las cajas de Petri con la nitrocelulosa se incubaron invertidas durante 2 hs a 37°C.

A continuación se retiraron los discos de nitrocelulosa cuidadosamente, y se los lavó durante 10 min con TBS. En el caso de requerir una réplica, se volvió a colocar un nuevo filtro con IPTG y se incubó durante otras 3 hs a 37°C.

* El IPTG actúa por desrepresión de la transcripción a partir del promotor P Lac de λ gt11 al unirse al represor lac, por consiguiente se produce la síntesis de las proteínas de fusión β -galactosidasa-ADNc de *T.cruzi*. Las proteínas de fusión liberadas por la lisis bacteriana son adsorbidas a la nitrocelulosa.

Detección de las proteínas con anticuerpos.

Las proteínas transferidas a nitrocelulosa se detectaron, en todos los casos, siguiendo las indicaciones de los fabricantes del kit Vecta Stain ABC (Vector Laboratories, Burlingame, CA, USA).

Bloqueo e incubación con el suero

Para evitar la unión inespecífica de los anticuerpos a la nitrocelulosa, el filtro se colocó en solución de bloqueo (SB: 2% glicina, 4% leche descremada (Molico), en TBS), durante 20 min como mínimo.

Se retiró la solución anterior y se incubó con el suero, previamente diluido 1:1000 – 1:2000, en SB conteniendo 0,05% de Tween 20, 5% de lisado de *E.coli* RY1089 lisogénica para kgt11 wt, 0,05% Tween 20, 1 mM PMSF, 0,01% azida.

La incubación se realizó a temp. amb. con agitación, por 3 hs o durante la noche, o a 4°C, durante la noche cuando se requirió recuperar la dilución del suero.

Los anticuerpos no unidos se eliminaron con 3 lavados de 5 min en TBS-TW.

Se agregó el segundo anticuerpo (anti-IgG humana total, biotinilado, preparado en cabra (kit ABC Vecta Stain, Vector Laboratories, Rurlingame, CA, USA)), diluido 1:200 en SB-TW y se incubó por 45 min.

Luego de 3 lavados de 5 min en TBS-TW, se colocó el complejo avidina-biotina-peroxidasa, (preparado 30 min antes como sugieren los fabricantes: 2 gotas de A (avidina DH) + 2 gotas de B (peroxidasa biotinilada) en 10 ml de TBS) y se incubó por 45 min más. El exceso de complejo se eliminó con 3 lavados de 3 min en TBS.

Revelado inmunoenzimático de la unión proteína de fusión-IgG de suero-anti IgG biotinilada-avidina-peroxidasa biotinilada.

Para el revelado, se agregaron 15 mg de 4-cloro-1-naftol (Merck, Darmstadt, Germany) disueltos en 5 ml de metanol frío y 8 μ l de H_2O_2 30 % a 20 ml de 50 mM Tris-HCl pH: 6.8. Esta solución se volcó inmediatamente sobre el filtro y se incubó en la oscuridad hasta la aparición de las señales (5 a 15 min).

La reacción se detuvo lavando el filtro con H_2O

Para la detección de proteínas utilizando anticuerpos inmuno-purificados se siguió el protocolo descripto sin diluir la solución de anticuerpos.

(1) Todas las incubaciones y lavados se efectuaron a temperatura ambiente y con agitación, adaptando los

volúmenes de cada reactivo al número y tamaño de los filtros, de modo de mantenerlos siempre cubiertos y sin superponerse entre sí.

3. HIBRIDIZACIONES RADIATIVAS

Condiciones de hibridización.

Los filtros fueron colocados dentro de bolsas de polietileno selladas. Por cada cm² de filtro Hybond N+ o nitrocelulosa se utilizó 0,05 ml de solución de hibridización. En casos de rastreo con filtros redondos se colocaron varios en una misma bolsa. Se usó 1 ml de solución de hibridización por filtro de 9 mm de diámetro, más 2 ml adicionales por bolsa; o bien 3 ml por filtro de 15 mm de diámetro más 5 ml adicionales por bolsa.

Generalmente, las hibridizaciones se llevaron a cabo en condiciones de alta rigurosidad según se detalla a continuación.

Hibridización con Formamida:

Se prehibridizó por dos hs a 42 °C en una solución conteniendo SSC 5x, formamida 50%, Denhardt's 5x, EDTA 5 mM, PO₄H₂Na 50 mM pH 6,5 y ADN de timo de ternero 250 pg/ml precalentado a 100°C durante 5 min.

En la experiencia de Northern blot, la solución de prehibridización incluyó Formamida 50%, Sulfato de Dextrano 10%, NaCl 1M, SDS 1%, Denhardt's 5x, ADN de timo de ternero 100 pg/ml precalentado a 100°C por 5 min.

La sonda (Act.Esp. 10⁷-10⁸ cpm/ug ADN) se agregó desnaturalizada (100°C por 5 min) a 0,5-1.10⁶ cpm/ml de solución de hibridización.

Se hibridizó a 42°C, con agitación lenta, durante la noche.

Hibridización rápida:

En algunas experiencias se usó la mezcla de hibridización rápida, *rapid hybridization system* de la firma Amersham (UK). La prehibridización se llevó a cabo a 65°C durante 15 min, luego se inyectó la sonda radioactiva, desnaturalizada a 100°C por 5 min. La hibridización se realizó a la misma temperatura durante 2 hs.

Condiciones de lavado:

Generalmente, se realizaron dos lavados a temp. amb. durante 30 min en 2x SSC/0,2% SDS.

Dos lavados a 67°C durante 30 min en 0,2x SSC/0,2% SDS. En casos donde se obtuvo una señal fuerte (>100 cps) se realizaron lavados adicionales en 0,1x SSC/0,2% SDS hasta 67°C.

Después de cada lavado los filtros se monitorearon con un contador Geiger hasta obtener una relación señal-ruido de fondo diferenciable.

Los filtros se expusieron sobre películas AGFA-GEVAERT Curix RP1 con pantalla intensificadora, a -70°C, por tiempo variable de acuerdo a la marca retenida.

4. MANIPULACIÓN DE FAGOS RECOMBINANTES

Purificación de un clon positivo.

Extracción de fagos con señal positiva.

Siguiendo las marcas del filtro, se ubicó la zona correspondiente al fago con señal positiva en el rastreo inmunológico. Es útil hacer un calco del filtro en papel transparente para luego, sobre un transiluminador ubicar las placas de lisis de la zona correspondiente a la señal positiva que se pueda extraer con una pipeta Pasteur de diámetro grueso, o con una pipeta automática cuyo cono se ha cortado para que tenga un diámetro mayor. El fragmento de agar con las placas de lisis, se aspira y se resuspende en 1 ml de SN con 30 ul de cloroformo, y se deja eluyendo por lo menos una noche antes del rastreo siguiente.

Rastros sucesivos: purificación de fagos con señal positiva.

Para purificar el fago recombinante positivo, se repite el rastreo inmunológico con los fagos extraídos de la zona correspondiente a la señal positiva, a una densidad de ufp por caja de Petri menor. Los fagos que a su vez produzcan una señal positiva se vuelven a rastrear, y el procedimiento se repite respetando las cantidades de ufp por caja de Petri que se indican a continuación. Cuando en dos rastreos sucesivos (generalmente el cuarto y el quinto), todos los fagos producen una señal positiva, se considera purificado, el fago puede ser amplificado para su posterior caracterización.

Placa grande 15 cm

Placa chica 9 cm

Amplificación de los fagos:

Para obtener una suspensión de fagos de aproximadamente 10^{10} ufp/ml, se infectaron 200 ul de bacterias *E.*

coli Y1090 preparadas como se detalla más adelante con 10 μ l de la dilución de fagos del último rastreo (aprox. 10^6 ufp/ml). Luego de 20 min a 37°C se mezclaron con 3 ml de top-agarosa (previamente fundida y mantenida a 50°C) y se sembraron en una placa de Petri (96 mm) de LB-agar-amp. Esta se colocó a 42°C. Al cabo de 4 hs se observó el crecimiento a confluencia de los fagos. Se agregaron 5 ml de SM por placa y se incubó durante toda la noche a 4°C. Al día siguiente se pasó la suspensión de fagos a un tubo de polipropileno. Los fagos remanentes se colectaron en 2 ml de SM que se agregaron a la preparación anterior. Se agregó cloroformo (3% final) y se dejó agitando suavemente durante 15 min a temp. amb.. Las bacterias no infectadas y los restos celulares se separaron por centrifugación a 3000g x 10 min a 4°C. Se colectó el sobrenadante, se agregó cloroformo (0,3% final) y se mantuvo a 4°C. El título del sobrenadante obtenido fue de aproximadamente 10^{10} 10^{12} ufp/ml.

Escisión in vivo de pBS SK⁻ de λ ZAPII

- Se usa la cepa JM109 que es RecA⁻.
- Mezclar 200 μ l de JM109 (DO₆₀₀=1) + 100 μ l λ ZAP (10^5 ufp) o 10 μ l λ ZAP (10^8 ufp) + 20 μ l Helper R408 ($1,5 \cdot 10^{11}$ ufp/ml). Colocar todo en un tubo falcon de 50 ml.
- Incubar 15 min. a 37°C
- Agregar 5 ml de 2x YT
- Continuar la incubación 4 Hs más a 37 °C con agitación.
- Incubar la mezcla a 70°C por 20 min para inactivar a λ ZAP y matar a las bacterias.
- Centrifugar a 1000g 5 min.
- Pasar el sobrenadante a un tubo estéril (contiene al fago f1 helper y a f1 con pBS)
- Mezclar 200 μ l de JM109 + 200 μ l de fago f1.
- Incubar 15 min a 37°C
- Plaquear sin diluir y una dilución de 10^{-2} para probar.
- Si el número de colonias es adecuado, picar y crecer toda la noche en LB o SOB a 37°C

5. PREPARACIONES DE ACIDOS NUCLEICOS

Preparación de ADN de fagos Lambda:

El ADN de fagos recombinantes se purificó utilizando el método siguiente:

- Se inocularon 10 ml de medio LB - Mg²⁺- maltosa. en tubos Falcon de 50 ml, con 200 μ l de células crecidas durante toda la noche en LB Mg²⁺ maltosa, y un taquito de agar conteniendo una placa de lisis del fago correspondiente.
- Para λ gt11 se utilizó la cepa Y1090, realizándose la infección a 42°C con agitación a 200 rpm

durante 5-6 hs.

- Para λ gt10 se utilizó la cepa LE392, realizándose la infección a 42°C con agitación de 200 rpm durante la noche.
- Como control de lisis se incluyó un tubo conteniendo células sin infectar. Concluido el tiempo necesario para la lisis bacteriana, evidenciada por la diferencia de D.O. entre los tubos infectados y el control de células sin infectar, se agregó 200 μ l de CHCl_3 y se incubó durante 5 - 10 min mas a 37 °C, con el objeto de lisar las bacterias remanentes.
- Se centrifugó durante 5 min a 4000 rpm, para eliminar los restos celulares y una alícuota del SN fue conservada a 4°C como stock de fagos (título aproximado 10^{10} ufp/ml).
- Se tomaron 8 ml de SN, se mezclaron con 8 ml de buffer TM, pasando todo el contenido a un tubo Corex de 30 ml. Se agregó 320 μ l de ADNasa I/ ARNasa A (1 mg/ml de cada una en TM).
- Se mezcló por inversión y se incubó 15-30 min a T.amb.
- Se concentraron luego los fagos por precipitación en PEG-NaCl, por agregado de 2 ml de NaCl 5M, 2,2 g de PEG 6000. Se disolvió completamente el PEG mezclando por inversión.
- Se dejó por 15 min en hielo
- Se centrifugó por 10 min a 12000 g a 4°C.
- Se descartó el SN, se resuspendió el pellet de fagos en 300 μ l de buffer TM. Se transfirió a un tubo Eppendorf. Se agregó 300 μ l de CHCl_3 , se mezcló bien y se centrifugó por 5 min para separar las fases.
- Se transfirió la fase acuosa superior a un tubo nuevo, cuidando de no remover la interfase de PEG. Se realizaron 2 nuevas extracciones.
- Se agregó a la fase acuosa 15 μ l de EDTA 0,5 M pH 8 y 30 μ l de NaCl 5 M.
- Se agregó 350 μ l de fenol equilibrado con Tris pH 7,6, se mezcló con vortex y centrifugó durante 5 min para separar las fases.
- Se repitió la extracción.
- A la fase acuosa se le agregó 875 μ l de etanol. Se incubó durante 20 min a -70 °C y se recogió el ADN por centrifugación durante 15 min.
- Se lavó el pellet con etanol 80 %. Se centrifugó por 2 min a 4°C, se descartó el etanol y se secó al vacío.

El ADN fue resuspendido en 50 μ l de buffer TE. El pellet bacteriano se resuspendió en 90 μ l 25 mM Tris-ClH pH8,0, 10 mM EDTA, 50 mM glucosa.

La suspensión de bacterias se transfirió a un tubo Eppendorf.

- Se agregó 10 μ l de lisozima 50 mg/ml disuelta en buffer Tris-EDTA-glucosa.
- Se mezcló suavemente por inversión y se dejó a temp. amb durante 5 min.
- Se agregó 200 μ l de NaOH 0,2 N - SDS 1% preparado en el momento.
- Se mezcló por inversión y se incubó en hielo por 15 min.
- Se centrifugó durante 10 min a 4°C en una microcentrifuga. Cuidadosamente se removió 350 μ l del SN y se transfirieron a un tubo Eppendorf.
- Se realizaron 2 extracciones con 1 vol de fenol-cloroformo isoamílico (25:24:1). Se centrifugó por

15 min a temp.amb.

- Se descartó el sobrenadante y se dejó drenar el exceso de líquido por inversión del tubo abierto sobre papel absorbente.
- El pellet se disolvió en 100 µl de TE.
- Se trató con 2,5 µl de ARNasa 10 mg/ml durante 10 min a temp.amb.
- Se agregó 80 µl de H₂O y 20 µl de NaAcO 3 M, pH 5,2.
- Se extrajo una vez con fenol-cloroformo, tomando los 180 µl superiores de la fase acuosa, a la cual se le hizo otra extracción con cloroformo, recuperando los 160 µl superiores de la nueva fase acuosa, que fue transferida a un nuevo tubo Eppendoff.
- Se agregó 2 vol de etanol absoluto, se mezcló por vortex y se dejó precipitar por 20 min a -70°C.
- Se recuperó el pellet por microcentrifugación y se lavó con etanol 70%. Se drenó el etanol remanente y se secó el pellet al vacío. El ADN fue resuspendido en 20 µl de Tris 10 mM pH 7,6. El rendimiento aproximado fue de 3 µg de ADN/ml de cultivo. Los plásmidos recombinantes preparados por este método resultaron adecuados como templates de secuenciación de ADN de doble cadena.

Extracción de ADN de Plásmidos:

En pequeña escala (Minipreparaciones).

Los plásmidos fueron preparados según dos métodos distintos, según el objetivo del experimento:

Met.A. para rastreo de plásmidos recombinantes.

Met. B. para mapeo de restricción y secuenciación de ADN dc.

Método A:

- A partir de una colonia aislada se inocularon 3 ml de medio LBampi y se incubaron con agitación a 37 °C hasta saturación.
- Se transfirió 1 ml de bacterias a un tubo Eppendorf y se cosechó por centrifugación a 4000 rpm por 5 min.
- Se descartó el sobrenadante y el pellet se resuspendió en 70 µl de buffer STET (Sacarosa 8 %, Tritón X-100 5 %, EDTA 50 mM, Tris.HCl pH 8,0 50 mM) empleando vortex.
- Se agregó lisozima (0,7 mg/ml final), mezclando con vortex.
- Se hirvió la suspensión durante 40 seg y se pasó a hielo.
- Se centrifugó durante 10 min en frío.
- Se descartó el precipitado con la ayuda de un escarbadiates o *tip* p20/p200.
- Se extrajo una vez con fenol-cloroformo-alcohol isoamilico (25:24:1).

- Se precipitó con 1 volumen de isopropanol. Se dejó a temp. amb. durante 15 min.
- Se centrifugó durante 5 min. a temp. amb.
- Se descartó el SN y se secó al vacío.
- Se retomó el pellet en 20 µl de buffer TE.
- Se usaron 3 µl para analizar plásmidos recombinantes por restricción agregando al buffer de digestión ARNasa libre de ADNasas, a una concentración final de 50 µg/ml.

Método B:

Se utilizó una modificación del método de Birnboim y Doly (1979) para secuenciar ADNdc.

- A partir de una colonia aislada se inocularon 5 ml de LB + 50 µg/ml de ampicilina en tubos de polipropileno (Falcon 2070) de 50 ml.
- Se incubaron con agitación durante la noche a 37 °C.
- Las bacterias fueron cosechadas por centrifugación a 4000 rpm por 15 min a temp. amb.
- Se agregó 16 ml de NaOH, 0,2N, 1% SDS (p/v); se mezcló por inversión y se dejó por 5 min a temp. amb.
- Se añadió 12 ml de KAc 0,5M pH 5,2 frío y se mezcló por inversión.
- Se dejó en hielo durante 10 min.
- Se centrifugó a 10000g durante 30 min a 4°C. Se recuperó el sobrenadante y se le agregó 0,6 vol de isopropanol, se mezcló y se dejó 15 min a temp. amb.
- Se centrifugó a 12000 g por 30 min, se descartó el sobrenadante y se resuspendió el pellet plasmídico en 7 ml de buffer TE.

Purificación de ADNccc por centrifugación en gradientes de CsCl-Bromuro de Etidio.

- Por cada ml de solución de ADN, se agregó 1 g de CsCl, mezclando suavemente hasta lograr la disolución de la sal.
- Se añadió 0,8 ml de una solución de Br. Et. (10 mg/ml en H₂O) por cada 10 ml de solución de ADN/CsCl, se mezcló bien y se estimó la densidad final de la solución pesando el tubo (aprox. 1,55 g/ml).
- Se centrifugó la solución a 8000 rpm por 5 min a temp. amb.
- Se descartaron los complejos de proteínas bacterianas formados, transvasando con cuidado la solución con ayuda de una pipeta Pasteur a un tubo de ultracentrifuga.
- Se ultracentrifugó en rotor 50 Ti a 36000 rpm, a 18°C durante 36 hs.
- La banda inferior correspondiente al ADN ccc fue extraída con pipeta Pasteur.
- El Br. Et. fue removido por extracción (7 veces) con butanol saturado en H₂O.
- La fase acuosa obtenida fue llevada a 8 ml con H₂O, agregando 0,8 ml de NaOAc 3M pH 5,2, más 17,6 ml de etanol y se dejó en tubo Corex de 30 ml durante 15 min a - 70 °C.
- El precipitado se recogió por centrifugación y se resuspendió en H₂O. El ADN se cuantificó a 260 nm y la calidad de la preparación se verificó en un minigel de agarosa.

- El rendimiento fue de 500 ug ADN plasmídico/ml de cultivo.

Preparación de ADN del fago M13.

Preparación de ADN de la forma replicativa del Fago M13.

La forma replicativa del Fago M13 (RF) es ADN de circular. Su preparación se llevó a cabo de la siguiente manera:

- Se utilizó un stock de fagos M13 de 10^{11} ufp/ml para infectar 2,5 ml de un cultivo de bacterias JM101 ($DO_{600} = 0,8$), se mezcló y se dejó durante 5 min a temp. amb.
- Las células infectadas fueron diluidas en 250 ml de LB precalentado a 37 °C.
- Se incubó por 5 hs a 37 °C con agitación vigorosa. Se cosecharon las células por centrifugación a 4000 xg, por 15 min a 4°C.
- El sobrenadante se removió y el pellet se resuspendió en buffer STE frío (0,1 M NaCl, 10 mM Tris. Cl pH 8,0, 1 mM EDTA pH 8,0).
- Se recuperaron las células lavadas por centrifugación a 4000 xg por 15 min a 4°C.
- Luego el ADN RF de M13 fue aislado por el método de lisis alcalina y purificado por centrifugación en gradiente de CsCl-Br.Et. según se detalló en la sección

Preparación de ADN simple cadena del Fago M13.

A. En pequeña escala:

- Se partió de una placa aislada (formada por inhibición de crecimiento bacteriano) de fagos M13 recombinantes. Se tomó la placa incluida en un bloque de agar, o bien se tomó una porción de la misma con un escarbiente o con una punta para micropipeta P-200.
- La placa fue transferida a un tubo de 2 ml de 2x YT conteniendo 10 ul de un cultivo saturado de células JM101.
- Se incubó durante 5-6 hs a 37 °C.
- Se transfirió a un tubo Eppendorf, se centrifugó a 5000 rpm por 10 min y el sobrenadante se transfirió a un tubo nuevo. Se volvió a centrifugar por 5 min.
- Se transfirió 1,2 ml del sobrenadante a un tubo nuevo al que se le agregó 200 ul de una solución 20% PEG 6000/2,5 M NaCl. Se mezcló por inversión y se precipitó durante 15 min a temp. amb.
- Se centrifugó durante 10 min, se descartó el sobrenadante, se centrifugó nuevamente para recoger el PEG remanente en las paredes del tubo, que se removió cuidadosamente con pipeta Pasteur afilada.
- El pellet se resuspendió en 200 ul de TE y se hizo una extracción con 100 ul de fenol: se mezcló con vortex por un min., se dejó reposar por 5 min, se volvió a mezclar y finalmente se separaron las fases por centrifugación.

- Se transfirió la fase superior cuidando de no tomar material de la interfase y se realizó una nueva extracción.
- Se extrajo una vez con 3 vol. de éter-H₂O y se precipitó la fase acuosa con NaAcO 0,3 M y 2,5 vol de etanol absoluto.
- Se dejó precipitando a -70°C durante 20 min, se centrifugó durante 15 min, descartando el sobrenadante.
- El pellet fue lavado con etanol 70%, secado al vacío y disuelto en 22 u1 de buffer TE.

Preparación de ADN total de *Trypanosoma cruzi*

El ADN genómico de *T. cruzi* se preparó siguiendo el protocolo descrito a continuación:

- Se resuspendió suavemente 1 g (peso húmedo) de parásitos en 20 vol de buffer 1. Se agregó proteinasa K (100 ug/ml final) y se incubó durante la noche a 37°C.
- Se extrajo 3 veces con fenol y 3 veces con cloroformo-isoamílico
- Se recuperó la fase acuosa.
- Se precipitó el ADN en 2,5 vol de Etanol absoluto y 0,1 vol de NaAcO 3M, pH 6.
- Se resuspendió en 10 ml de TE y se llevó a una concentración de 0,83g/ml de CsCl más 0,01% de BrEt. Se realizó una ultracentrifugación en Sorvall OTD 65B con rotor de ángulo fijo 50ti a 40000 rpm, durante 36 Hs a 18°C. Se recuperó la banda de ADN lineal.
- Se extrajo el BrEt 7 veces con butanol 2. Se precipitó en presencia de NaAcO 0,3 M y 2,5 vol de Etanol absoluto durante 15 min a -70°C, y se recuperó el ADN por centrifugación durante 20 min a 10000 rpm a temp. amb.
- Se lavó 2 veces con Etanol 70% a T.amb y se secó al vacío.
- Se guardó disuelto en TE a una concentración de 500 ng/μl.

Preparación de ARN total de *Trypanosoma cruzi*

Método de extracción con isotiocianato de guanidinio (GTC) en colchón de Cloruro de Cesio.

Para garantizar la calidad de las preparaciones de ARN se utilizó material de vidrio esterilizado por horneado durante varias horas a 200°C, H₂O bidestilada tratada con 0,1% DEPC, material de plástico nuevo autoclavado 3 veces y guantes nuevos.

- Se resuspendió el pellet de parásitos (se utilizó entre 500 mg y 1 g de pellet peso húmedo) en 2,5 ml de GTC.
- Se pasó 30-60 veces por jeringa G-23 para romper el ADN.
- Se mezcló por vortex intensamente durante 5 min.
- Se colocó cuidadosamente el homogenato sobre 1,5 ml de CsCl 6M/EDTA 0,1 M, constituyén

dose 2 fases; el CsCl es la inferior y el homogenato la superior.

- Se centrifugó a 35000 rpm en ultracentrifuga Sorvall OTD 65B durante 18 hs en SW 56 a 18°C. Al final de la corrida se visualizó un pellet de ARN.
- Se extrajo la fase superior más 500 ul de la solución de CsCl con pipeta Pasteur de modo de no alterar el pellet de ARN se lavaron las paredes del tubo con 500 ul de GTC. Se volvió a extraer la fase superior. Este procedimiento se repitió 2 veces más
- Finalmente, se descartó todo el sobrenadante de CsCl.
- El pellet fue resuspendido en 200 ul de GTC con pipetman P-1000, punta azul hasta lograr su resuspensión completa (unas 50 veces).
- El ARN fue transvasado a un tubo Eppendorf, y se agregó 3 vol de Etanol. Se precipitó a -70°C durante 1 hora.
- Se centrifugó a 12000 xg durante 30 min a 4°C.
- El pellet se lavó con 1 ml de Etanol 70% evitando resuspenderlo. Se centrifugó a 12000 xg por 5 min. Se realizó otro lavado.
- El pellet se secó al vacío durante 5 min y se resuspendió en 100 ul de TE.
- El ARN resultó adecuado para experiencias de Northern blot y clonado por PCR. Para la construcción de bibliotecas de ADN copia se utilizó ARN poli A+

6. CLONADO Y SUBCLONADO EN VECTORES. TRANSFORMACIÓN

CLONADO POR ARN-PCR

Aproximadamente 2 ug de ARN total de *T. cruzi* fueron incubados con retrotranscriptasa utilizando hexámeros al azar en la siguientes condiciones de reacción:

buffer PCR Cetus II 10x	2ul
MgCl ₂ 25 mM	4ul (5mM)
RNAsin (Promega)	1 ul (2 U)
dNTPs	8 ul (4 mM)
Hexámeros al azar	1 ul (2,5 uM)
AMVtranscriptasa reversa (Promega)	2 ul (40 U)
ARN total (preincubado a 70°C, 2')	2 ul (2 ug)

La reacción fue incubada a 25 °C por 10 min, a 42°C por 40 min, y luego a 95°C por 5 min.

El volumen total de la reacción fue transferido a un tubo conteniendo 80 ul de :

PCR bufferI	8ul
MgCl ₂ 25 mM	4ul (2mM)
primer 3' específico	1ul(0.15uM)
primer 5' SL	1,6ul(0.15uM)
Taq polimerasa(Promega)	1ul(2.5U)
H ₂ O csp 80 ul	64,4ul

Se agregaron 50 ul de aceite mineral (Sigma). La reacción se llevó a cabo en un aparato programable de la marca HYBAID modelo IHBlO1. Se utilizó el siguiente programa de amplificación:

CICLO	PASO	TEMP.(°C)	TIEMPO	LOOPS
1	1.01	94°C	5'	1
	1.02	52°C	1'	
2	2.01	72°C	1'	40
	2.02	94°C		
	2.03	52°C	1'	
3	3.01	72°C	10'	1

Luego de concluido el programa de amplificación se sembró 1/10 de la reacción en un minigel de agarosa 2 % para verificar la amplificación.

Para hacer a estos insertos susceptibles de subclonado, fue necesario rellenar los extremos para generar productos de extremos romos (*blunt end*).

Los productos de la reacción de PCR fueron incubados a 37°C por 15 min con 2 U de Klenow (no fue necesario agregar dNTPs porque la reacción de PCR llevó dNTPs en exceso).

Se hizo una extracción con fenol-cloroformo 1:1 y el producto se precipitó con 1/10 vol de NaAcO 3M pH 6 y 2 vol de Etanol absoluto. Se dejó 20 min a -70°C, se centrifugó durante 10 min a 11000xg, se lavó con Etanol 70% y se secó el pellet.

El producto de PCR fue purificado por elución de un gel de agarosa de bajo punto de fusión.

Dado que los productos de PCR no llevan grupo fosfato en sus extremos 5' fue necesario fosforilar los extremos del producto purificado por la acción de la polinucleótido Kinasa.

Subclonado en vector M13:

protocolo:

Se tomo 1 ug de ADN ds λ gt11 cortado con EcoR1 que representa 10 ng de inserto/ul y se diluyo 1/10.

RELACION	VECTOR	INSERTO	BUFFER	ATP	AGUAdd	LIGASA T4
	10 ng		10X	10 mM		3.75 U
1:1	1 ul	0.5 ul	1 ul	1 ul	5.5 ul	1 ul
1:5	1 ul	3.0 ul	1 ul	1 ul	3.0 ul	1 ul
	1 ul	1 ul	1 ul	1 ul	6.0 ul	1 ul

El volumen final es de 10 ul. La ligasa de T4 (7.5 U/ul)(PHARMACIA).

La ligacion se llevo a cabo a 16°C durante toda la noche.

Subclonado de los productos de PCR. generación de "blunt-end DNA". reacción de ligación.

Los productos de la reacción de PCR fueron incubados 15 a 37°C con 2U de klenow para completar los extremos (los dNTPs no se agregan porque están en exceso).

Luego el producto se precipitó y preparó para fosforilar los extremos.

Se resuspendió de la siguiente manera:

- 2 ul 5X PNK buffer
- 1 ul 0.1 M DTT
- 2 ul 100 uM ATP
- 4 ul agua
- 1 ul T4 polinucleótido kinasa

Se incubo 30 min a 37°C

Se trato la enzima a 65°C unos minutos.

Las reacciones de ligación se llevaron a cabo como se describo anteriormente, excepto que se uso el vector M13 mp10 cortado con SmaI (AMERSHAM). Además se usaron 50 ng de vector y 7.5 U de ligasa de T4. La reacción se hizo a 22°C toda la noche.

Preparación de células competentes para transformación.

En este trabajo, se utilizo la cepa E.coli DH5αF. Esta cepa tiene ciertas ventajas sobre otros hospedadores de M13:

- Pueda ser mantenida en medio rico (LB o YT) sin perdida del episoma P, porque a diferencia de la serie JM no depende de la presión de selección dada por el marcador proAB en el F. Ello obligaba a mantener a la cepa en medio mínimo.

- El genotipo incluye el marcador RecA1 que asegura la estabilidad de los insertos clonados, la serie JM es rec+.
- Es restricción negativa (rk-) y modificación positiva (mk+), así que cualquier secuencia puede ser fácilmente clonada.
- Tiene una mayor eficiencia de transformación ($> 10^7$ ufp/ml).
- Esta cepa también tiene un background genético lac- para permitir α -complementación con el polipéptido codificado por M13 serie mp.

La cepa DH5 α F se mantiene en presencia de kanamicina.

- Se estrián colonias frescas la noche anterior y se crecen a 37°C en placas LB.
- Se pican 4-5 colonias y se ponen en un eppendorf conteniendo SOB. Se crecen 1 hora a 37°C.
- Pasar el inóculo a 50 ml de SOB en erlenmeyer de 500 a 1000 ml.
- Incubar con agitación aireada a 37°C hasta DO₅₅₀ = 0.45-0.55 ($8 \cdot 10^7$ células/ml).
- Pasar el contenido a dos tubos corex de 30 ml. Centrifugar a 4000 rpm 10 a 4°C.
- Resuspender en 1/3 de volumen de TFB. Incubar 15 min. en hielo.

TFB = KCl 1000 mM
 MnCl₂ 45 mM
 CaCl₂ 10 mM
 H₂CoCl₂ 3 mM
 K-Mes 10 mM

Mes = 0.5 M Mes pH=6.5 equilibrado con KOH

- Centrifugar a 4000 rpm 20 a 4°C.
- Resuspender en 1/12 de volumen de TFB.
- Agregar DnD 3.5% final. Mezclar e incubar en hielo por 10 min

DnD = DTT 1 M
 DMSO 90%
 AcoK 10 mM

- Agregar DnD 7% final. Incubar por otros 20 min en hielo.
- Las células ya están listas para transformar en alícuotas de 200 μ l

Con este método denominado "Standard High Efficiency Transformation" se obtiene una eficiencia de transformación entre $5 \cdot 10^6$ - $5 \cdot 10^7$ ufp/ μ g de RF.

Transformación de los productos de ligación.

Paralelamente a la preparación de competentes se hace un inóculo en 3 ml de LB de E.coli JM101 a saturación, para ser usadas en la infección.

- Se mezclan 5 ul de ligación con 200 ul de DH5 α F competentes.
- Incubar 30 min a 0°C.
- Dar un shock de calor 2 min a 42°C .
- Incubar en hielo durante otros 5 min.
- Agregar 200 ul de JM101 saturado
 - 40 ul de X-gal 2%
 - 20 ul de IPTG 100 mM
- Transferir a soft-agar fundido y mantenido a 50°C .
- Plaquear sobre LB-agar.
- Incubar toda la noche a 37°C.
- Los subclones positivos se identificaron por hibridación.

Minipreparación de ADN simple cadena de M13

- A partir de una placa blanca, picar con escarbadiente y sumergir en un tubo con 2 ml de LB + 15 ul de bacterias JM101.
- Dejar creciendo 5 Hs. a 37°C con agitación.
- Centrifugar a 3500 rpm 10 a temperatura ambiente.
- Tomar 1.5 ml del sobrenadante y pasar a un Eppendorf.
- Centrifugar a 3500 rpm 10 min.
- Tomar 1.2 ml del sobrenadante sin tomar bacterias y pasar a otro Eppendorf (guardar el resto con 100 ul de glicerol a -20°C).
- Agregar 150 ul de PEG 20%-AcoNH₄ 7.5 M. Homogeneizar por inversión y dejar a temperatura ambiente 15 min.
- Microcentrifugar 10 min . Aquí debe verse el pellet, sino no seguir.
- Descartar el sobrenadante con pipeta pasteur. Remover bien los restos de PEG.
- Resuspender en 100 ul de TE.
- Agregar 50 ul de fenol equilibrado. Vórtex 1.
- Centrifugar 5 min. Pasar fase acuosa a otro eppendorf. Repetir.
- Hacer 2 extracciones con éter saturado de 200 ul c/u. Evaporar el éter a 37°C unos minutos.
- Agregar 2.5 volúmenes de etanol abs. y 1/2 volumen de AcoNH₄ 7.5 M. Poner 15 min a -70°C.
- Microcentrifugar 15 min. Sacar sobrenadante. Lavar con etanol 75% y secar al vacío.
- Resuspender en 22 ul de TE.

Complementariedad por retraso de hebra para determinar la orientación del inserto subclonado.

El método consiste en verificar el retardo de la migración en geles de agarosa del ADNsc de un fago recombinante M13 cuando al ADNsc de un inserto puede formar un híbrido parcial con otro fago M13 que lleva su cadena complementaria (por inserción en la otra orientación). Así se identifican pares de fagos M13 que contienen las dos orientaciones del inserto.

Se parte de la minipreparación de ADN simple cadena y se juntan de a pares las preparaciones.

- 5 ul ADNsc A (100-200 ng)
- 5 ul ADNsc B (100-200 ng)
- 1 ul NaCl 2.5 M
- Desnaturalizar 30 a 60°C.
- Hibridar enfriando lentamente en baño seco a 60°C hasta temperatura ambiente.
- Agregar 2 ul de loading buffer.
- Correr un minigel de agarosa 0.7%.
- Seleccionar aquel par que haya sufrido retraso en la corrida.

7. SECUENCIACIÓN

secuenciación por la técnica de Sanger usando el kit Sequenasa (U.S.B.)

Las reacciones de secuenciación por Sanger se llevaron a cabo con el kit SEQUENASE VERSION 2.0 de U.S.B..

La secuenciación incluye 3 pasos:

- ◆ Anillamiento del cebador de la reacción de polimerización.
- ◆ Reacción de marcado
- ◆ Reacción de terminación

Se usaron las siguientes soluciones:

Sequenase buffer 5X= 200 mM Tris-HCl pH=7.5

100 mM MgCl₂

250 mM NaCl

Labeling mix (dGTP) 5X= 7.5 uM dGTP

7.5 uM dCTP

7.5 uM dTTP

ddG Termination mix (dGTP)= 80 uM dGTP, 80 uM dATP, 80 uM dCTP, 80 uM dTTP, 5 uM ddGTP, 50 mM NaCl.

ddA Termination mix (dGTP)= 80 uM dGTP, 80 uM dATP, 80 uM
dCTP, 80 uM dTTP, 8 uM ddATP, 50
mM NaCl.

ddT Termination mix (dGTP)= 80 uM dGTP, 80 uM dATP, 80 uM
dCTP, 80 uM dTTP, 8 uM ddTTP. 50
mM NaCl.

ddC Termination mix (dGTP)= 80 uM dGTP, 80 uM dATP, 80 uM
dCTP, 80 uM dTTP, 8 uM ddCTP, 50
mM NaCl.

Enzyme dilution buffer= 10 mM Tris-HCl pH=7.5
5.0 mM DTT
0.5 mg/ml BSA.

Mn buffer= 0.15 M isocitrato de sodio
0.1 M MnCl₂.

Stop solution= 95% formamida
20 mM EDTA
0.05% azul da bromofenol
0.05% xilen cianol

Protocolo Standard :

Se utilizó generalmente una masa de ADNsc de 500 ng. El volumen de la solución de ADNsc no debe exceder los 7 ul.

A) Mezcla de anillado:

ADN 7 ul
buffer sequenasa 2 ul
"primer" 1 ul
total 10 ul

Calentar a 65°C 2 min en un baño seco y dejar enfriar lentamente hasta una temperatura menor a 35°C.
Luego colocar en hielo.

Mientras se enfría se preparan 4 tubos conteniendo c/u 2.5 ul de la mezcla de terminación (ddA, ddC, ddG, ddT) y se rotulan correctamente. En este caso se usaron eppendorf de diferentes colores para cada terminación.

- Diluir la mezcla de marcación 1:5 generalmente 2 ul mezcla + 8 ul agua.
- Diluir la Sequenasa version 2.0 en cantidad suficiente para todos los templados en el buffer de dilución 1:8.
- Precalentar en un baño de agua a 37°C los 4 tubos de terminación preparados.

B) Reacción de marcado:

- A la mezcla de anillado agregar:

DTT 0.1 M 1 ul
mezcla de marcación 1:5 2 ul
[35S]dATP 0.5 ul
Sequanasa 1:8 2 ul

Mezclar e incubar a temperatura ambiente durante 3 min . Este paso es estricto pues la extensión de la marca no permitir leer las bases más cercanas al iniciador.

C) Reacción de terminación:

- Transferir 3.5 ul de la reacción de marcado a cada tubo de terminación (ddA, ddC, ddG, ddT) y continuar la incubación a 37°C por 5 min.
- Parar las reacciones por agregado de 4 ul a cada tubo de STOP SOLUTION.
- Calentar las muestras a 75°C por 2 min antes de cargar en el gel.
- Se siembran 3 ul de cada letra.

En los casos de ensayo con Mn Buffer se realiza igual pero se agrega 1 ul de Mn Buffer durante la reacción de marcado.

Armado y corrida de un gel de poliacrilamida desnaturazizante para secuenciación.

Los geles desnaturizantes de poliacrilamida para secuenciación corridos fueron de 6 % que permiten leer hasta 400 bases desde el inicio de la reacción.

Se preparó una solución stock de acrilamida 40% de la siguiente forma:

Acrilamida 380 g
N,N -metilenbiscrilamida 20 g

Agua para 100 ml

Se disolvió la solución por calentamiento a 37°C. La solución para los geles de secuencia se armó en 100 ml en erlenmeyer de la siguiente forma:

42 g urea (7 M final)

15 ml de acrilamida 40 % (6 % final)

TBE 1X hasta 100 ml

Se disuelve la urea. Poner en hielo hasta ser volcado, justo antes de volcar en los vidrios se agrega:

100 ul TEMED

280 ul persulfato de amonio 10 % para iniciar la polimerización.

Se utilizaron vidrios de cuarzo de 30 X 40 cm. Los vidrios se limpiaron bien y luego se siliconaron con un algodón embebido en silic-glass. Se usaron espaciadores de 0.4 mm y se sellaron los vidrios con cinta. Luego se volcó el gel con una inclinación de 45° continuamente para no interrumpir el flujo y evitar burbujas de aire. Normalmente la polimerización lleva alrededor de 30 min, pero si el gel no polimeriza en una hora debe ser descartado.

Antes de que polimerice se agrega el peine en posición invertida en la parte superior y se deja toda la noche a temperatura ambiente. Al otro día se monta sobre la cuba en posición vertical.

Se coloca el peine introduciéndolo en el gel hasta pinchar la acrilamida y se agrega buffer TBE 1X en la parte superior e inferior de la cuba. Se precorre al gel a 60 Watts durante 30 min antes de sembrar para que los vidrios adquieran una temperatura homogénea.

Sembrar 3 ul de cada letra. En un gel 6 %, el azul de bromofenol corre como 25 pb y el xilen cianol 100 pb. Como regla general, el segundo depósito se sembró 30 min después de salir el azul. La corrida del gel se paró cuando salió el azul del segundo depósito.

Se desmontó de la cuba y se separó un vidrio (el gel queda adherido al otro). Se fijó durante 15 min en una mezcla 10 % metanol + 10 % ácido acético. Se lavó con abundante agua. Luego se colocó sobre el gel un Whatman 3 MM al cual debe quedar adherido el gel.

Se cubrió con Saran-Wrap y se secó durante 3 Hs. en una secadora de vacío Hoeffler modelo SE 1160. Finalmente se expuso con pantalla intensificadora a -70°C durante 2 días.

8. TÉCNICAS DE MANEJO DE ACIDOS NUCLEICOS

Southern blots

En todos los casos el ADN genómico se diluyó a 100 ng/ul y se usaron de 3 a 5 U de enzima de restricción por ug de ADN. Se digirieron 10 ug para cada caso. Las digestiones se realizaron a 37°C toda noche. Se

corrieron geles de agarosa 0.9% a 5 volts/cm durante toda la noche. Luego se fotografió bajo luz U.V.

Posteriormente se siguieron dos procedimientos alternativos:

A) El gel se trató 30 min en una solución desnaturalizante 0.4N NaOH, 0.6M NaCl y se neutralizó dos veces de 30 en 0.5N Tris-HCl pH 7.6, 1.5M NaCl. Se transfirió a membranas de nylon Hybond N+ (cargadas positivamente) en una solución 20X SSC durante toda la noche. Se fijó en una solución alcalina de 0.4N NaOH por 20 min.

B) El gel no se pre-trató como en A y se transfirió a membranas de nylon Hybond N+.y se fijó en una sola etapa en una solución alcalina conteniendo 0.4N NaOH

La hibridación con sondas radioactivas se hizo como se describió anteriormente. Se expuso a -70°C con pantalla intensificadora durante tiempos variables de 2 a 7 días según el caso.

Dot blot de ARN

El dot blot se realizó usando diluciones decrecientes al medio de ARN total de *T.cruzi* de la siguiente forma:

100 µg 50 µg 25 µg 12.5 µg 6.25 µg

Se usó también ARN sintético de TcP1 como control en diluciones al medio de la siguiente manera:

10 ng 5 ng 2.5 ng 1.25 ng 0.625 ng

Los Dots de ARN se depositaron sobre un filtro de nylon Hybond N+ (Amersham, USA) usando un aparato de Dot blot comercial y una bomba de vacío.

Las sondas utilizadas se marcaron a una actividad específica similar como se determinó usando un contador de centelleo líquido. Las sondas tenían entre 3-6.108 cpm/µg.

Luego de hibridación, la marca asociada a los Dot blots se determinó por autorradiografía y por conteo en centelleo líquido de cada punto del Dot blot. Se graficaron los datos de marca (cpm) vs. Concentración del Dot (µg) y se obtuvieron rectas aplicando regresión lineal. La pendiente de cada recta (cpm/µg) se utilizó para graficar la concentración relativa de cada ARNm estudiado.

9. TRANSCRIPCIÓN Y TRADUCCIÓN IN VITRO

Transcripción Run Off

1. Preparación del templado

Se utiliza la técnica Birnboim y Doly para preparar plásmidos a partir de 10 ml de cultivo.

- Se resuspende en 1 ml de STE y se pasa a un eppendorf.
- luego de preparación, se resuspende en 500 µl de TE/RNAsa
- Se incuba a 37°C durante 15 min.
- Se agrega un volumen de fenol, se mezcla y separan las fases.
- Se eliminan los restos de fenol con 2 vol. de ETER:agua.
- Se agregan 500 µl 1.6 M NaCl + 13 % PEG 8000. Mezclar.
- Centrifugar en microcentrífuga 5 min.
- Resuspender en 100 µl de TE
- Agregar 10 µl de 10X proteínasa K buffer y 10 µl de 5% SDS.
- Agregar 100 µg/ml de proteínasa K
- Incubar 1 hora a 37°C.
- Agregar un volumen de fenol, mezclar y separar la fases.
- Resuspender en 50 µl TE RNAsa free.
- Volver a precipitar con la solución de PEG8000/NaCl.
- Resuspender en 50 µl de TE RNAsa free
- Cortar el templado con EcoRI , 5 U/ µg de templado, en 50 µl.
- Digerir con proteínasa K nuevamente.
- Agregar un volumen de fenol, mezclar y separar las fases.
- Precipitar con 2 vol de EtOH Abs. y ½ vol de AcoNH₄ 7.5 M
- Secar el pellet
- Resuspender en 15 µl de TE RNAsa free.

2. Transcripción a partir de los promotores T7 y SP6.

Sin Capping

- | | | |
|---|-------------------|-------|
| ■ | ADN templado 2 µg | 4 µl. |
| ■ | DTT 100 mM | 5 µl |

- 5X buffer de transcripción 10 μ l
- rNTPs 10 mM 2.5 μ l c/u
- BSA 10 mg/ml 0.5 μ l
- RNAsin (o HPRI) 40 U/ μ l 1.5 μ l
- T7 RNA polimerasa 15 U/ μ l 2 μ l
- Incubar 1 hora a 37°C

Con Capping

Se utiliza un protocolo similar, a excepción de agregarse una concentración 10 veces menor de rGTP y el análogo de CAP GpppG 5 mM 5 μ l.

- Se agrega RQDNAsa 1 μ l (1mg/ml)
- Incubar 15 min a 37 °C
- Agregar un volumen de fenol, mezclar y separar las fases.
- Precipitar con 12 μ l de AcoNH₄ 10M y 150 μ l de EtOH abs.
- Resuspender en 50 μ l de agua RNAsa free.
- Realizar una segunda precipitación igual.
- Resuspender en 10-20 μ l de agua RNAsa free.
- Se corre una alícuota en un gel de formaldehído agarosa.
 - 1.5 g agarosa
 - 20 ml 5X MOPS
 - 50 ml agua DEPC
 - a 50°C agregar 16 ml de formaldehído deinizado.

TRADUCCIÓN IN VITRO

1. Puesta a punto de las condiciones

Se utilizó el kit de reticulocitos de conejo de Amersham, USA N90Z. Este kit contiene un pool de aminoácidos no marcados.

La reacción se lleva a cabo en un volumen de 50 μ l y se utiliza [35S] MET (1000 ci/ mmol) a una concentración final e 2 μ ci/ μ l. La reacción standard usa 80 % de lisado en la mezcla.

Los factores que afectan la traducción son los siguientes:

- Concentración de Mg⁺⁺ y K⁺, para decrecer la concentración en la mezcla de reacción se puede diluir el lisado hasta un 40%. Esto varía para cada ARN a traducir en particular.
- Concetración de ARN, a mayor concetración se aumenta la cantidad de producto marcado.

- La concentración del aminoácido marcado.

Se probaron las condiciones para la concentración óptima de ARN y de concentración de lisado. Se utilizó ARN total de *T. cruzi* para las pruebas.

- Lisado 20 μ l (40 %) o 40 μ l (80 %)
- ARN 1 a 5 μ l
- [35S]Met 7 μ l
- Agua 22 - 17 μ l o 3 - 0 μ l
- Volumen total 50 μ l
- Se incuba 1 hora a 30°C
- Se destruyen los complejos [35S]Met-tRNA con 0.3 NaOH durante 15 min a 37°C

La prueba de traducción usando ARN poliA+ o ARN total mostró que con ARN poliA+ se llega rápidamente a una saturación de producto marcado con aproximadamente 200 ng de ARN. En cambio con ARN total, el producto fue aumentando correlativamente en una gama de concentración de 2 a 10 μ g. Obteniéndose con 10 μ g de ARN total, la misma cantidad de producto que con 100 ng de ARN poliA+

- La marca asociada a cada reacción se midió en un contador de centelleo.
- Para disminuir el background al medir, los filtros se lavaron en TCA 25% Caseína hidrolizada 2%
- De allí se toman 10 μ l para contar en centelleo líquido con omnifluor
- Omnifluor: 5 L tolueno, 4g PPO, 0.2g POPOP
- Finalmente, se trata el producto con RNAsa A durante 10 min.
- Se calienta en cracking buffer 5-8 min a 100°C
- Se toman 15 μ l (dependiendo de la marca) para analizar por PAGE-SDS 14 %
- Luego de la corrida se realizó fluorografía (para mejorar la detección de la marca en autorradio grafía.)
- Se realizan 2 lavados en DMSO 30 min c/u
- Incubar 3 horas en PPO 20 % en DMSO
- Lavar con abundante agua (importante!)
- Secar el gel durante aproximadamente 1 hora.
- Exponer toda la noche.

3. Inmunoprecipitación de la traducción in vitro

- Tomar 1 mg de proteína A en 1 ml de buffer NaHCO₃ 0.1 M (pH=10)/NaCl 0.5 M/B-

mercaptoetanol 1mM

- 100 mg CnBr Sepharosa 4B se hidratan en 20 ml de 1 mM Hcl 10 min.
- Se filtra el HCl y se lava con 50 ml de 0.1 M NaHCO₃/0.5 M NaCl (pH=8.5)
- Pasar el gel seco a un recipiente con la solución de la proteína A. Incubar a 4°C 18 horas con agitación.
- Colectar el sobrenadante y medir a DO280, por diferencia sacar la cantidad unida.
- El gel se satura con Tris-Hcl 0.1 M lavando y concentrando 3 veces.
- Lavar el gel en buffer acetato 0.1 M, 0.5 M NaCl y con tris-Hcl 0.1 M, NaCl 0.5 M 3 veces alternativamente.
- Lavar con PBS
- Tomar 5 µl de la traducción in vitro
- Agregar 10 µl de suero y 60 µl de buffer I = 1% nonadict P40/ 0.1% SDS/ 0.5% deoxicolato de sodio/ 0.02% azida sódica/100 µg/ml PMSF.
- Incubar durante 3 horas a 4 °C con agitación
- Se agregan 60 µl de proteína A-Sepharosa en buffer I y se incuba 30 min a 4°C
- Se recoge el complejo terciario por centrifugación 20 min.
- Lavar el pellet 5 veces en 1.5 ml de 0.1% nonadict P40/100 µg/ml PMSF en PBS
- Centrifugar 20 seg. Y descartar el sobrenadante
- Solubilizar en 20 µl de craking buffer por calentamiento a 100°C.
- Sembrar

10. TÉCNICAS ESPECIALES DE PCR

Inter SIRE y SIRE-Bubble PCR

La técnica de SIRE-PCR es particularmente poderosa porque puede ser aplicada sin conocimiento previo del segmento genómico a analizar. Sin embargo, tiene limitaciones. SIRE-PCR sucederá únicamente entre pares de secuencias SIRE que estén a una distancia y orientación apropiada una con respecto a la otra. La utilidad de SIRE-PCR está además restringida por la aparente distribución asimétrica de las secuencias repetidas dentro de los genomas. Esta organización particular, que da como resultado regiones ricas y pobres en contenido del elemento, lleva a una distribución no uniforme de productos de PCR, dejando algunas zonas del genoma subrepresentadas.

SIRE-Bubble PCR permite superar estas limitaciones con una estrategia simple. Utiliza un “linker” específico en un extremo, independiente de la secuencia a analizar y un cebador específico de SIRE en el otro (el detalle de la técnica se puede ver en el esquema de la figura...). De esta forma se genera una distribución más uniforme de productos que representan mejor al genoma.

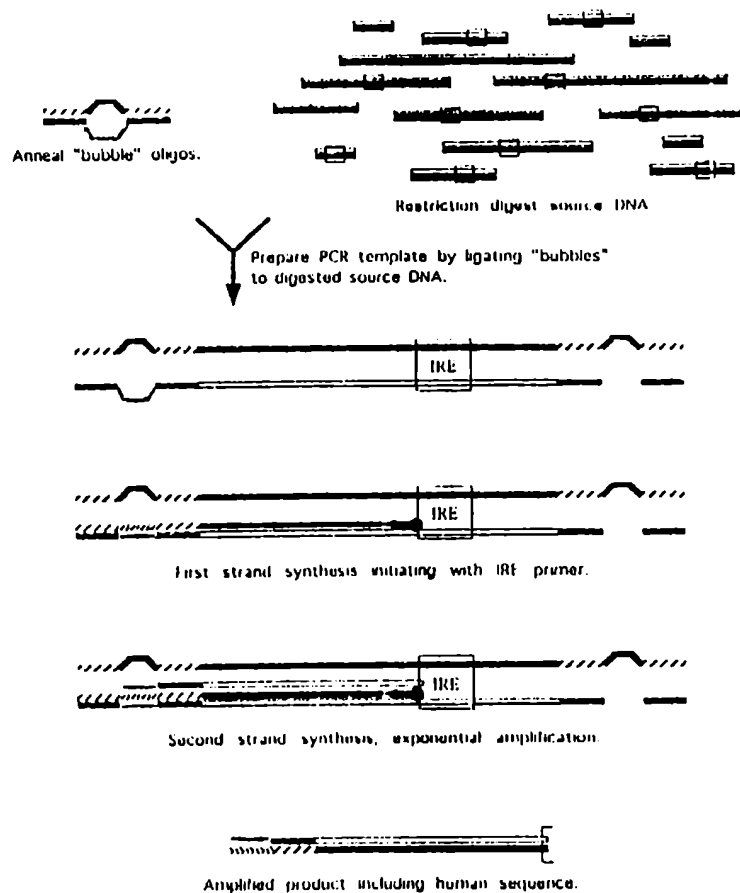


FIG. 1. Construction of IRE-bubble PCR templates and the IRE-bubble PCR amplification scheme.

Los cebadores utilizados fueron los siguientes, ubicados dentro de la secuencia de SIRE:

SIRE-1: 5'-GGAGGACCCAAAGTCGTCC-3'

SIRE-2: 5'-GATCTTGGGAGAGCTGGCTA-3'

SIRE-3: 5'-CTTACGAAGTGGCAGACTTT'-3'

HA-1: 5'-GCCAGCTCTCCCACGATCATTC-3'

HA-2: 5'-GGATCCGGCCCCAGCTGCCCGGAGGATTATT-3'

También se utilizaron los cebadores siguientes para la reacción de Bubble:

Bubble top: 5'-

GAAGGAGAGGACGCTGTCTGTCTGAAGGTAAACGGACGAGAGAAGGGAGAG-3'

Bubble bottom: 5'-

CTCTCCCTTCTGCGGCCCGCAGTTCGTCAACATAGCATTTCTGTCCTCTCCTTC-3'.

Bubble primer: 5'-GCGGCCGCAGTTCGTCAACATAGCATTTCT-3'

1) **INTER-SIRE PCR**

ADN genómico o de YAC	50-100 ng (2 µl)
Buffer 10X	10 µl
MgCl ₂ 25 mM	10 µl
dNTPs 10 mM c/u	2 µl
DMSO	10 µl
Cebador correspondiente	100 ng (2 µl)
agua	58 µl
Taq polimerasa 5 U/µl	0.5 µl
Total	100 µl

Temperatura de anillamiento = 58°C

Numero total de ciclos = 30

2) **SIRE Bubble PCR**

2.1- Preparacion de los anclajes “Bubble”.

- Mezclar el oligo Bubble top y Bubble bottom a 3 pmol/µl en 2X SSC/100 mM Tris en un volumen total de 333 µl.
- desnaturalizar la mezcla a 100°C por 10 minutos y dejar enfriar lentamente a temperatura ambiente.

2.2- Construcción de los templados de PCR SIRE-Bubble.

- Digerir 1 µg de ADN genómico con RsaI o AluI en un volumen total de 120 µl.
- Incubar la reaccion a 37°C durante 4 Hs.
- Calentar a 65°C durante 20 minutos para inactivar la enzima.
- Mezclar 60 µl (0.5 µg) con 12 pmoles (4 µl) de anclajes “Bubble” en un volumen total de 100 µl.
- Agregar 15-20 U de T4 ADN Ligasa e incubar a temperatura ambiente toda la noche.
- Diluir la mezcla de reaccion a 1.5 ng/µl con ¼ X TE.

2.3- Condiciones de PCR

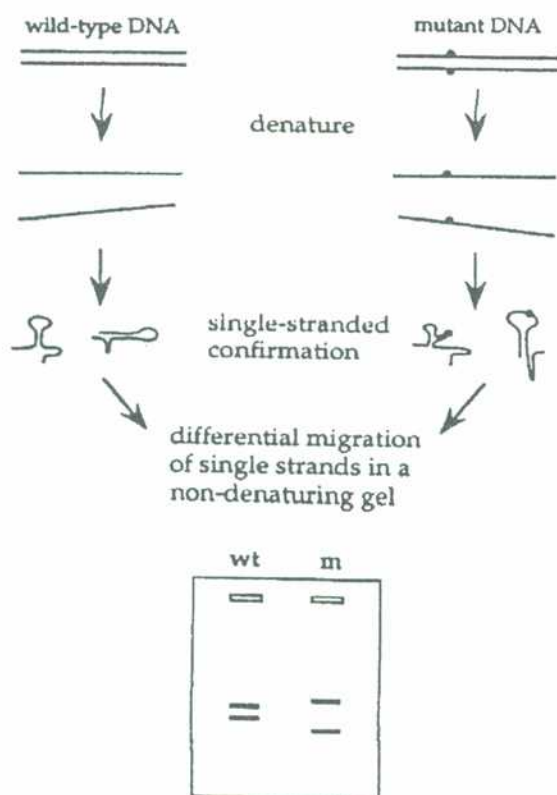
Se inició con un procedimiento “Hot Start” en 50 µl de reacción

Templado de ADN	10 ng (2 µl)
Buffer 10X	5 µl
MgCl ₂ 25 mM	5 µl
dNTPs 10 mM c/u	1 µl
Bu-primer	100 pmol (2 µl)
HA-1 primer	100 pmol (1 µl)
Agua	30 µl
Taq polimerasa 5 U/µl	0.5 µl
Total	50 µl

Temperatura de anillamiento = 60°C

Número total de ciclos = 32

ANÁLISIS SSCA-PCR DE FRAGMENTOS GENÓMICOS



La movilidad de una partícula en un gel es sensible tanto a su tamaño como a su forma en condiciones no desnaturilizantes. El ADN de cadena simple tiene una estructura secundaria que es determinada por interacciones intramoleculares, y conbsecuentemente por su secuencia. En el análisis SSCA (antes conocido como SSCP, Single Strand Conformational Polimorphism), una secuencia mutada es detectada como un cambio en la movilidad en geles de poliacrilamida causada por una alteración en su estructura secuendaria.

Hasta el presente no es posible predecir el cambio en la movilidad electroforética inducida por una mutación, pero el gran poder de resolución de los geles de poliacrilamida pueden distinguir la mayoría de los cambios causados por diferencias en la secuencia como ser una **única** sustitución en una molécula de cientos de bases de longitud (en general hasta 400-500 bases).

ADN Templado*	Variante clonada H6.4	Variante clonada H1.8	Variante clonada H1.5	A D N genómico cepa G	A D N genómico cepa Tu- lahuen
DNA**	2 µl	2 µl	2 µl	2 µl	2µl
Buffer 10X	5 µl	5 µl	5 µl	5 µl	5 µl
MgCl ₂ 25 mM	5 µl	5 µl	5 µl	5 µl	5 µl
DMSO 10% final	-	-	-	5 µl	5 µl
Primer S-1 100 ng/µl	1 µl	1 µl	1 µl	1 µl	1 µl
Primer S-2 100 ng/µl	1 µl	1 µl	1 µl	1 µl	1 µl
dNTPs 10mM c/u	1 µl	1 µl	1 µl	1 µl	1 µl
[α ³² P]dCTP 3000 Ci/mmol	0.5 µl	0.5 µl	0.5 µl	0.5 µl	0.5 µl
Water	hasta 50 µl	hasta 50 µl	hasta 50 µl	hasta 50 µl	Hasta 50 µl
Taq polimerasa 5 U/µl	0.5 µl	0.5 µl	0.5 µl	0.5 µl	0.5 µl

1. Amplificación radiactiva por PCR de las variantes TcP2β.

*DNA genómico debe ser precalentado a 100°C durante 5 min.

** 100 ng de DNA genómico y 1 ng de plásmido conteniendo la secuencia clonada de la variante TcP2β.

Temperatura de anillamiento = 52°C, 1 minuto

Temperatura de desnaturalización = 94°C, 1 minuto

Número total de ciclos = 32

Tamaño esperado del fragmento = 100 bp. .

2. Preparación de la solución de gel

Los geles de poliacrilamida SSCA pueden ser preparados de dos maneras: sin glicerol o con la adición de 10% de glicerol.

	Sin glicerol	Con 10% de glicerol
Acrilamida 40 % (1.5 % Bis-acrilamida)	9 ml	9 ml
TBE 10X	3 ml	3 ml
Glicerol	-	6 ml

Agua	48 ml	42 ml
TEMED	50 μ l	50 μ l
APS 10 %	250 μ l	250 μ l

El gel resultante es 6% total de acrilamida y 5% C, donde C es el grado de cross-linking (una tasa del porcentaje de concentración de N,N'-methylenebisacrylamide con respecto a la concentración total de acrilamida monómero. Un gel con bajo %C es más blando, tiene un tamaño de poro mayor y parece ser más sensible a la conformación.

3. Sembrado y corrida del gel

- ♦ Mezclar entre 1 y 3 μ l de muestra + 9 μ l de solución de stop Sequenase™.
- ♦ Calentar las muestras en agua hirviendo por 5 min. antes de sembrar.
- ♦ Cargar 3 μ l de muestra en el gel ($\sim 10^3$ - 10^4 cpm/ μ l).
- ♦ Correr el gel a 2 Watts durante 7 horas a toda la noche.
- ♦ Transferir el gel a whatmann 3 MM y exponer.

Buffer de corrida : 0.5 X TBE. (importante!)

Hay cuatro condiciones diferentes de corrida:

1. Gel sin glicerol a temperatura ambiente.
2. Gel sin glicerol a 4°C
3. Gel con glicerol a temperatura ambiente
4. Gel con glicerol a 4°C

La detección de mutaciones depende de los cambios conformacionales de la molécula de simple cadena inducida por la mutación, y consecuentemente, sensible a los cambios del ambiente físico en el gel, por ejemplo, temperatura, concentración de iones y solventes.



CARACTERIZACION INICIAL DE PROTEINAS RIBOSOMALES P DE *T.cruzi*

1.1- Revisión histórica

En el momento de comenzar el presente trabajo de tesis en 1992, el laboratorio estudiaba antígeno recombinante de *T.cruzi* que presentaba homología con las proteínas ribosomales P de humanos, al mismo se lo denominó JL5. Su secuencia parcial fue publicada por primera vez en 1989. Su importancia estaba dada por el alto nivel de anticuerpos contra este antígeno que poseía el suero de enfermos chagásicos crónicos.

El análisis de los aspectos inmunológicos de la reacción anti-JL5 fue motivo de varios trabajos y tesis (Gabriela Levitus, tesis doctoral, UBA, 1992; Dan Kaplan, tesis doctoral, UBA, 1995).

Parte del análisis genético de la proteína codificada por JL5 fue abordado por otro miembro del laboratorio (Alejandro Schijman, tesis doctoral, UBA, 1992), y finalmente se obtuvo la secuencia codificante completa del ARNm de esta proteína publicada en 1990.

El objetivo de la presente tesis fue caracterizar el sistema de las proteínas ribosomales P genéticamente en detalle y ampliar los conocimientos sobre las mismas mas allá de los que se tenían hasta la actualidad.

A medida que fuimos avanzando sobre el estudio, sospechamos que este sistema de proteínas podía ser un excelente modelo para el análisis de aspectos genéticos generales de este parásito. Cuanto más nos sumergimos en la profundidad de su genética, mas sorpresas nos deparaba y mas nos entusiasmaba.

En 1995 llegamos al final de este camino esperando haber cumplido nuestro objetivo mínimo de arrojar un poco mas de luz sobre la oscura genética este

Capítulo 1

Objetivos:

- Rastreo y caracterización de fagos recombinantes que codifican porciones de proteínas ribosomales P de *T.cruzi*.

tripanosomatido. Esperamos también haber sembrado un camino fértil para que mis compañeros continúen con este fascinante sistema genético de *T. cruzi*.

1.2- Obtención de ADNc que codifican para proteínas ribosomales P de *T. cruzi*.

Es conocido, ver introducción, que la familia de proteínas ribosomales P en otros organismos esta constituida por tres proteínas distintas que comparten un extremo C-terminal en común (P1, P2 y P0). En los eucariontes inferiores las proteínas P1 y P2 se presentan en dos variantes cada una (y). La proteína codificada por JL5 (que al momento de ser publicada recibió el nombre de P-JL5) presenta homología con

SONDAS

A: JL5



B: CH3



C: Lyt 6



D: G13



Disposición de los fagos

51)	S18	JL5	wt
○	○	○	○
JL52	CH3	Lyt 6	G13
○	○	○	○

las proteínas de tipo P2 (9, 14).

Con el objetivo de obtener secuencias de los ARNm de todos los antígenos ribosomales P, se rastrearon dos bibliotecas de ADNc de expresión del parásito construidas en el laboratorio en el vector GT11, una de la cepa Tulahuen (epimastigotes) y otra de la cepa RA (tripomastigotes sanguíneos).

Por rastreo de estas bibliotecas con anticuerpos anti-JL5 inmunopurificados de un suero chagásico, se obtuvieron varios fagos recombinantes. Además de la secuencia completa de los ARNm de la proteína P-JL5 se esperaba encontrar algunos que codificaran otras proteínas ribosomales P,

dado que los anticuerpos anti-JL5 reconocen la zona C-terminal común a todas las proteínas P. El mismo suero de un paciente con alto nivel de anticuerpos contra proteínas P aisló otro recombinante que se llamo JL52 (E. Mesri, tesis doctoral,

Figura 1-

Panel de fagos recombinantes hibridizados con las sondas radiactivas indicadas en la figura

UBA, 1990; A. Schijman, tesis doctoral, UBA, 1992). Anticuerpos inmunopurificados contra este el mismo, reconocían una proteína ribosomal del parásito de 38 Kd. Sin embargo, JL5 y JL52 no presentan reacción cruzada (A. Schijman, tesis doctoral, UBA, 1992).

En un primer análisis, todos los recombinantes aislados usando anticuerpos inmunopurificados contra estos dos antígenos recombinantes fueron probados por el ensayo de placas de lisis.

Figura 2.

Secuencia contenida en el fago G13. Se recuadra la zona codificante y se subraya el codon stop.

```

AGAAGGAAGA AGAGGAGGAG GACGATGACA TGGGCTTTGG TCTGTTTGAC 50
TAAGTGCAGC GGCATGCAAC GTGCCGGAGC AGAAAGGGAC TCGTCACTG 100
CTCTAGAACA AGGAAGAACA TACTGGCCGC CGTATTTGGG GTGTGCAAGA 150
CTGGAAGGCA TCTTGCAGTT CGGCAACTTT TTTTGATTCT GTTTTTTCTA 200
TTTATGAATG AATTTAAAAA AAAAAAAAAA 230
  
```

Esto permitió clasificarlos en dos grupos: 1) aquellos que reaccionaban con anticuerpos anti- JL5 y 2) aquellos que reaccionaban con anticuerpos anti- JL52.

Usando una sonda radiactiva JL5 en un ensayo de placas de lisis en el cual se ordenaron todos los fagos aislados (figura 1), se observó que no todos los fagos que reaccionaban con el anticuerpo anti-JL5 reaccionaban con la sonda de ADN (ver los fagos G-13 y Lyt 6 por ejemplo).

Luego, las sondas marcadas G13 y Lyt 6, cuyos ADNc codifican productos reconocidos por anticuerpos anti-JL5 pero no reaccionaban con fagos positivos para la sonda JL5, tampoco reaccionaban entre si (figura 1). Posiblemente estos fagos codificaban otras dos proteínas ribosomales P distintas a JL5. Por lo tanto, G13 y Lyt 6 se seleccionaron para posterior caracterización.

Entre los fagos que se agrupaban con JL5, solo el fago 5D fue seleccionado porque presentaba un tamaño de inserto que no era el esperado para una proteína de tipo P2.

Los fagos CH3 y JL52 eran serológicamente cross-reactivos, sin embargo sus ADN presentaron débil reacción cruzada en este ensayo. Como ninguno reaccionaba con los anticuerpos anti-JL5 no se continuó su análisis (ver A. Schijman, tesis doctoral, UBA, 1992).

Análisis del recombinante G-13

El fago G13 contenía un inserto de 220 pb con una zona 3' y una cola poli A. La zona codificante comprendía solo 15 aminoácidos (figuras 2 y 5).

La región 3' no codificante era distinta de JL5, pero la secuencia de aminoácidos de la zona codificante mostraba solo que la región C-terminal era igual a la de la proteína JL5.

Con el objetivo de clonar y secuenciar el ARNm completo de G13 para verificar si se trata de una proteína P nueva, se rastreó una biblioteca de ADNc de la cepa RA con la sonda radioactiva G13.

Entre 20000 fagos recombinantes se aisló uno que se denominó MAT1 y se purificó en sucesivos rastreos. El tamaño de su inserto es de 320 pb, según se calculó luego de la amplificación por PCR. Esto permitía suponer que si bien el nuevo ADNc no codificaba el ARNm completo extendía la secuencia hacia el

Figura 3

Secuencia contenida en el fago MAT1. Se recuadra la zona codificante y se subraya el codon stop.

GAATTCCTGC	AGCCGCTCCT	GCTGCTGCTG	CTGGCGGTGC	CGCCGCCCTT	50
GCCGCCGCTG	ACGCCAAGAA	GGAAGAAGAG	GAGGAGGACG	ATGACATGGG	100
CTTTGGTCTG	TTTGAC	<u>TAAG</u>	TGCAGCGGCA	TGCAACGTGC	CGGAGCAGAA 150
AGGGACTGCG	TCACTGCGCA	CTTGAACAAG	GAAAAACATA	CTGGCCGCCG	200
TATTTGGGGT	GTGCATGACT	GGAAGGCATC	TTGCAGTTTCG	GCAACTTTTT	250
TTTGATTCTG	TTTTTTCTAT	TTATGAATGA	ATTAAAAAAA	AAAAAAAAAA	300
AAAAAAAAAA	AAAAGGAATT	C	321		

extremo 5'.

El producto de PCR se fosforiló y se subclonó en el fago M13mp10 digerido con SmaI. Luego de la transformación se purificó el ADN de varios fagos.

Las cadenas complementarias se secuenciaron por la técnica de Sanger.

La secuencia de MAT1 contenía un larga región poli A (32 bases) y una secuencia de 320 bases. La secuencia es 60 bases más larga que G13, extendiéndola 20 aminoácidos en dirección N-terminal (figuras 3 y 5).

La secuencia aminoacídica deducida permite observar la zona "bisagra" rica en alaninas y confirma la presencia de los residuos lisina, ausentes en JL5 y característicos de otras proteínas P eucariontes en el extremo C-terminal (figura 5).

Análisis del recombinante Lyt 6

El fago Lyt 6 era una proteína ribosomal P por su reactividad inmunológica, pero seguramente era otra proteína de la familia por su reactividad con sondas de ADN.

El fago recombinante contenía un inserto pequeño de ~200 pb. Se secuenció por la técnica de Sanger, previo subclonado en M13 (figura 4).

Figura 4 -

Secuencia
contenida en el
fago Lyt6. Se
recuadra la zona
codificante y se
subraya el codon
stop.

```
GAATTCCCCG CGAAGAAGGA AGAGGAAGAG GAAGACGACG ATATGGGCTT 50
TGGTCTGTTT GACTAGGCGC AGAAGCCCCT TTCTTGAAGG ATTTACACAC 100
AAGTGGCATG GTTTTGTTTT TTGTGTTTTG TTTTCTCAAA TATTTATTCA 150
```

Presenta una cola poli A de 25 bases y un marco abierto de lectura que codifica para 18 aminoácidos interrumpido en 5' por el sitio de clonado. La secuencia 3' no codificante tiene ~110 bases y no presenta homología con ninguna de las otras dos proteínas P clonadas. Los 18 aminoácidos obtenidos son idénticos a los de MAT1, confirmando su identidad P, pero hay diferencias nucleotídicas que no cambian la secuencia de aminoácidos. De esta manera se confirma que P-G13 y P-Lyt6 son el producto de distintos genes y no de copias de un mismo gen. No se puede asegurar,

Figura 5 -

Comparación de
las secuencias
proteicas
contenidas en los
fagos G13, Lyt6,
MAT1 y JL5

G13	KEEEEEEDDDMGFGLFD

Lyt 6	AKKEEEEEEDDDMGFGLFD

MAT 1	AAAP--AAAA-GGAAAPAAADAKKEEEEEEDDDMGFGLFD
	* * * * * * * * * * * * * * * * * *
JL5	ASAPTAAAAASSGAAAPAAA-A---EEEEEDDDMGFGLFD

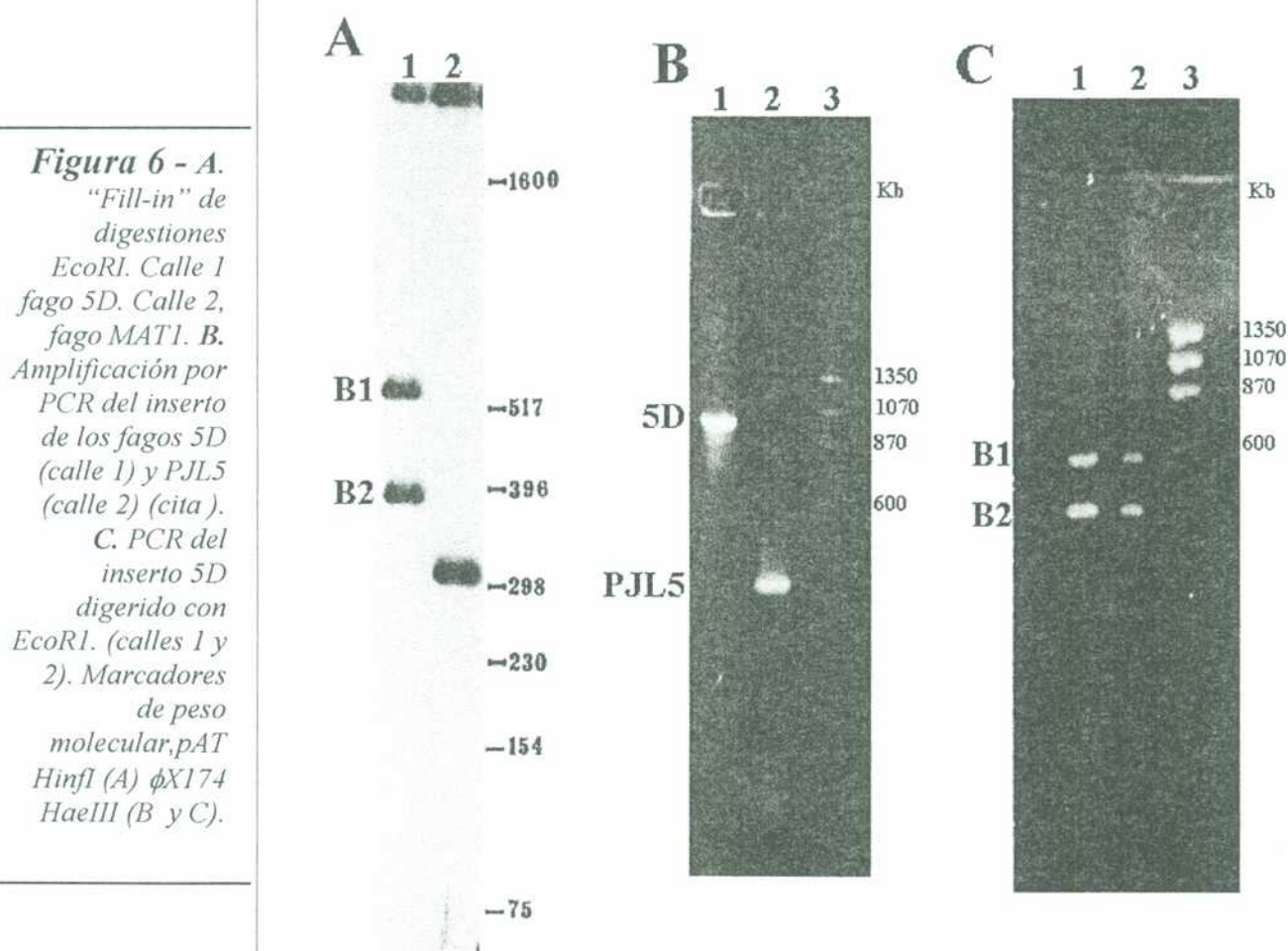
sin embargo, si codifican la misma proteína; más secuencia hacia 5' haría falta para determinar esto (figura 5).

Análisis del recombinante 5D

El fago 5D se relaciona inmunológica y radiactivamente con JL5. Fue obtenido de

una biblioteca de ADNc construida con un cebador sintético de la región C-terminal de JL5 (A. Schijman, tesis doctoral, UBA, 1992). Al determinar el tamaño del inserto por la técnica de "Fill-in" se obtuvieron dos bandas: una de 550 pb, B1 y una menor, de igual intensidad, de 400 pb, B2 (figura 6.A).

Por su tamaño, B2 parecía corresponder al ADNc completo de PJL5, de allí la sorpresa que causó la presencia de B1. En un principio se pensó que el fago 5D estaba impurificado con otro recombinante.



Entonces se plaqueó y rastreó radiactivamente a muy baja densidad (10 ufp/placa), pero todas las playas de lisis volvieron a dar señal positiva indicando que el fago estaba puro.

Era probable entonces que 5D contuviera a B1 y a B2 como un único inserto con un sitio EcoRI interno.

Para comprobarlo, el inserto se amplificó por PCR usando cebadores adyacentes al sitio de clonado. Si B1 y B2 estaban en el mismo fago se debía amplificar una única banda que fuera la suma de B1+B2, si por el contrario eran dos fagos que copurificaban se debía repetir el patrón del "fill-in". El resultado se observa en la figura 6. Se detectó una sola banda de 940 pb que podía ser la suma de B1 y B2 (figura 6.B). La comparación con PJL5 muestra la diferencia que existe entre los tamaños de ADNc.

Si realmente era la suma de dos bandas, el producto de PCR debía cortarse con EcoR1 y reproducir el patrón del "fill-in". Como se muestra en la parte C de la figura 6, el corte genera bandas de 560 pb (B1+cebador 1) y 420 pb (B2+cebador 2). Se confirmaba así que B1 y B2 están contenidas en el fago 5D.

Sin embargo, el tamaño del inserto era muy grande si tenemos en cuenta que debía codificar para PJL5.

Las dos bandas se purificaron por separado, se subclonaron en M13 y se secuenciaron.

El inserto B2 contenía 414 bases con una zona 5' no codificante de 162 bases y un marco abierto de lectura de 321 bases. La secuencia codificaba para el mensajero completo de otra copia PJL5 y lleva las últimas 22 bases del SL en el extremo 5' (figura 7).

El inserto B1 contiene 569 bases y adyacente a un sitio de clonado EcoR1 se encuentra la secuencia del oligonucleótido de JL5 usado para la construcción de la biblioteca (subrayado en la figura 7). Esto explicaba el origen del ADNc; aun así en ningún marco de lectura se encontraba la secuencia de PJL5. Sin embargo, hay un marco abierto de lectura de 117 aminoácidos que se inicia con un ATG en la posición 149 y termina con un codón stop en la posición 500, dejando la secuencia del oligonucleótido en la zona 3' no codificante. Los otros dos marcos son interrumpidos por codones stop.

Posteriormente, otros tres fagos recombinantes independientes obtenidos llevaban una secuencia idéntica a B1, aunque en los tres casos mas corta. Su posición se indica en la figura 7.

Evidentemente este oligonucleótido era capaz de hibridar (no necesariamente en su

Figura 7 -

A. Secuencia de la banda B2 de 5D. Recuadrada se marca la secuencia parcial del SL.

B. Secuencia de la banda B1 de 5D. Subrayada se señala la secuencia que presenta homología con el cebador específico P de *T. cruzi* usado para construir la biblioteca. Las flechas indican el comienzo de otras secuencias obtenidas que tienen homología con la banda B2.

A

ACGTTTCTGTACTATATTGATAATTTTCAGGATCCGGCCAGCTGCCCGGAGGATTATTCTGCTTTATAAGTTGTCTTG


TCAG	ATG	AAG	TAC	CTT	GCC	GCT	TAT	GCT	CTT	GTT	GGC	CTC	TCA	GGG	GGG	ACC	CCC	TCA
	M	K	Y	L	A	A	Y	A	L	V	G	L	S	G	G	T	P	S
AAG	TCT	GCA	GTT	GAG	GCC	GTA	TTG	AAG	GCT	GCC	GGC	GTT	CCT	GTC	GAT	CCT	TCT	CGC
K	S	A	V	E	A	V	L	K	A	A	G	V	P	V	D	P	S	R
GTG	GAT	GCA	CTT	TTT	GCC	GAG	TTT	GCA	GGC	AAG	GAC	CTT	GAT	ACG	GTA	TGC	ACT	GAA
V	D	A	L	F	A	E	F	A	G	K	D	L	D	T	V	C	T	E
GGC	AAA	TCG	AAG	CTT	GTA	GGA	GGG	GCT	ACA	CGT	CCC	AAT	GCT	GCC	ACC	GCT	TCT	GCA
G	K	S	K	L	V	G	G	A	T	R	P	N	A	A	T	A	S	A
CCC	ACT	GCC	GCT	GCC	GCC	GCC	TCT	AGT	GGT	GCC	GCC	GCG	CCT	GCT	GCA	GCT	GCG	GAA
P	T	A	A	A	A	A	S	S	G	A	A	A	P	A	A	A	A	E
GAG	GAG	GAG	GAC	GAT	GAC	ATG	GGA	TTT	GGT	CTT	TTT	GAC	TAA					
E	E	E	D	D	D	M	G	F	G	L	F	D	-					

B


GAATTCGCG ATGGCCGTGT AGTGGACTGC CATGGCGTTG ACGGGAGCGG GGGATTAGGG 60


TTCGATTCCG GAGAGGGAGC CTGAGAAATA GCTACCACTT CTACGGAGGG CAGCAGGCGC 120

GCAAATTGCC CAATGTCAAA AAAAAAGAT GAGGCAGCGA AAAGAAATAG AGCCGACAGT 180

GCTTTTGCAT TGTCGTTTTT AATGGGGGAT  G800 ATCCAAAATC GAGTAACAAT 240

TGGAGGACAA GTCTGGTGCC AGCACCGCGT AATTCAGCT CCAAAAGCGT ATATTAATGC 300

TGTTGCTGTT AAAGGGTTTC TAGTTGAATT GAGGGCTCT  2M AAGGCGCAAT GGTTTAGTCC 360

CATCCACTTC GGATTGGTGA CCCATGCCCT TGTGGTCCGT  P02 GAACAGACAT TCAGAAACAA 420

AAAACACGGG AGTGGTACCT TTTCTCTGAT TATCGCATGT CATGCATGCC AGAGGGCGCC 480

CGTGATTTTT TACTGTGACT AAAAAAGTGT GACCAAAGCA GTCATTCGAC TTGAATTAGA 540

AAGCATGGGA TTTGGTCTTT TTGGAATTC 569

totalidad) con otro mensajero del parásito y se clonó en esta biblioteca.

La secuencia de B1 se comparó en el banco de datos de secuencias EBI (European Bioinformatics Institute, Oxford, UK).

La comparación de la banda B1 dio el siguiente resultado:

Se obtuvo un 98.5 % de homología con la subunidad pequeña del ARN ribosomal de *T. cruzi* (ver figura 8.)

Dio además 89.6 % de homología con la misma subunidad de ARN ribosomal de *Leishmania donovani*; 89.4% con *Crithidia fasciculata*; 87.4% con *Trypanosoma brucei*.

CUTOFF : 300 KTUP : 4
TARGET : GenBank database RELEASE : R74.0
GROUP NAME : Invertebrate
2 TRBSSRRNA >T.CRUZI SMALL SUBUNIT RIBOSOMAL RNA GE 99.5 548 1132 2154
99.5% identity in 548 bp overlap. initial score: 1132 optimized score: 2154

Figura 8 -
Comparación de la banda B1 de 5D en el banco de datos del EBI. Se muestra la homología obtenida con la subunidad menor del ARN ribosomal de *T. cruzi*. Se realizó la comparación con CUTOFF = 300 y KTUP = 4.

	10	20	30	40	50	60
5DBANDA1.SEQ	GAATTCGCGCATGGCCGTGTAGTGGACTGCCATGGCGTTGACGGGAGCGGGGGATTAGGG					
TRBSSRRNA	TCAGCCAGCGATGGCCGTGTAGTGGACTGCCATGGCGTTGACGGGAGCGGGGGATTAGGG					
	350	360	370	380	390	400
5DBANDA1.SEQ	TTCGATTCCGGAGAGGGAGCCTGAGAAATAGCTACCACCTCTACGGAGGGCAGCAGGCGC					
TRBSSRRNA	TTCGATTCCGGAGAGGGAGCCTGAGAAATAGCTACCACCTCTACGGAGGGCAGCAGGCGC					
	410	420	430	440	450	460
5DBANDA1.SEQ	GCAAATTGCCCAATGTCAAAAAAAAAAAGATGAGGCAGCGAAAAAGAAATAGAGCCGACAGT					
TRBSSRRNA	GCAAATTGCCCAATGTCAAAAAAAAAAAGATGAGGCAGCGAAAAAGAAATAGAGCCGACAGT					
	470	480	490	500	510	520
5DBANDA1.SEQ	GCTTTTGCATTGTCGTTTTCAATGGGGGATATTTAAACCCATCCAAATCGAGTAACAAT					
TRBSSRRNA	GCTTTTGCATTGTCGTTTTCAATGGGGGATATTTAAACCCATCCAAATCGAGTAACAAT					
	530	540	550	560	570	580
5DBANDA1.SEQ	TGGAGGACAAGTCTGGTGCCAGCA-CCGC-GTAATTCCAGCTCCAAAAGCGTATATTAAT					
TRBSSRRNA	TGGAGGACAAGTCTGGTGCCAGCACCCGCGGTAATTCCAGCTCCAAAAGCGTATATTAAT					
	590	600	610	620	630	640
5DBANDA1.SEQ	GCTGTTGCTGTTAAAGGGTTCGTAGTTGAATTGAGGGCCTCTAAGGC GCAATGGTTTAGT					
TRBSSRRNA	GCTGTTGCTGTTAAAGGGTTCGTAGTTGAATTGAGGGCCTCTAAGGC GCAATGGTTTAGT					
	650	660	670	680	690	700
5DBANDA1.SEQ	CCCATCCACTTCGGATTGGTGACCCATGCCCTTGTGGTCCGTGAACAGACATTTCAGAAAC					
TRBSSRRNA	CCCATCCACTTCGGATTGGTGACCCATGCCCTTGTGGTCCGTGAACAGACATTTCAGAAAC					
	710	720	730	740	750	760
5DBANDA1.SEQ	AAAAAACACGGGAGTGGTACCTTTTCTCTGATTATCGCATGTCATGCATGCCAGAGGGCG					
TRBSSRRNA	AAAAAACACGGGAGTGGTACCTTTTCTCTGATTATCGCATGTCATGCATGCCAGAGGGCG					
	770	780	790	800	810	820
5DBANDA1.SEQ	CCCGTGATTTTTTACTGTGACTAAAAAGTGTGACCAAAGCAGTCATTTCGACTTGAATTA					
TRBSSRRNA	CCCGTGATTTTTTACTGTGACTAAAAAGTGTGACCAAAGCAGTCATTTCGACTTGAATTA					
	830	840	850	860	870	880
5DBANDA1.SEQ	GAAAGCATGGGATTGGTCTTTTTTGAATTC					
TRBSSRRNA	GAAAGCATGGGATAACAAAGGAGCAGCCTCT					
	890	900	910			

OBTENCION Y ANALISIS DE SECUENCIAS DE ARNm COMPLETAS DE LAS PROTEINAS RIBOSOMALES P.

2.1- Puesta a punto y obtención de secuencias de ARNm de proteínas P completas usando la técnica de ARN-PCR en *T.cruzi*.

Una vez obtenidas secuencias parciales de las nuevas proteínas P, el siguiente objetivo fue obtener sus secuencias completas. Las secuencias parciales de los ADNc G13(MAT1) y Lyt6 permitían distinguir si se trataba de dos genes que codificaban para la misma proteína o correspondían a dos de proteínas P diferentes. Para llevar a cabo este objetivo, se puso a punto una técnica de PCR partiendo de ARN total como templado, tal como se describe en materiales y métodos.

La idea del experimento se basa en que todos los ARNm de *T.cruzi* comienzan con una secuencia líder de 39 bases (SL). Así, si se tiene información sobre la secuencia 3' de un ARNm de interés, se puede diseñar un oligonucleótido sintético correspondiente a esa zona para usar como cebador 3' en una reacción de PCR.

Como la secuencia 5' de ese mensajero no es conocida se puede usar como cebador de esa zona un oligonucleótido que corresponda a la secuencia del SL. Este cebador se apareara con todos los ARNm, pero la especificidad de la reacción está dada por el cebador 3' específico de la secuencia que se quiere completar. Esta técnica puede ser considerada como una variación de la técnica de PCR anclado (anchored PCR) o RACE (Rapid Amplification of cDNA Ends) (61), ya que la secuencia de SL puede actuar como un anclado natural en la amplificación de mensajeros de tripanosoma.

Para poner a punto este protocolo se preparó ARN total de epimastigotes de *T.cruzi* de la cepa RA con un rendimiento total de 3 mg de ARN en excelente estado a partir de 1 g de parásitos.

Objetivos: El objetivo de este capítulo fue obtener las secuencias completas de las proteínas ribosomales P, analizarlas y clasificarlas.

El diseño de los cebadores se hizo de la siguiente manera:

A) Para el cebador 5' (oligonucleótido basado en SL) se usaron las primeras 20 bases de la secuencia de 39 (OLIGO MX)

5'- AACGCTATTATTGATACAGT- 3' T_m= 52C

B) Para el cebador 3' que correspondía al recombinante G13 se usaron las 18 bases río abajo del codón de terminación de la traducción. Este oligonucleótido es antisentido (OLIGO G13ext) (ver figura 2)

5'- CGTTGCATGCCGCTGCAC -3' T_m= 60C

C) Para el cebador 3' del fago recombinante Lyt6 se usaron 23 bases ubicadas a 25 nucleótidos río abajo del codón de terminación (OLIGO L6 antisentido) (ver figura 4).

5'- CCATGCCACTTGTGTGTAAATCC- 3' T_m= 68C

El método tiene la ventaja que permite usar ARN total y evita la purificación de ARN poli A+, simplificando el método y abreviando los tiempos.

El primer paso del método es la síntesis de la primera cadena de ADNc a partir del ARN total. Esta síntesis se hizo usando hexámeros al azar, también puede realizarse usando oligo dT como cebador. Sin embargo el uso de hexámeros al azar aumenta considerablemente la eficiencia de síntesis.

El primer ciclo de la reacción de PCR es selectivo de las moléculas de interés al unirse el cebador 3' específico.

Luego de la síntesis de ADNc a 42°C, una fracción de la muestra se utiliza para la reacción de PCR agregando los cebadores adecuados, es decir OLIGO MX y OLIGO

G13ext/OLIGO L6 según sea el caso.

Como controles de especificidad, se realizaron reacciones paralelas en las cuales no se incluyó ARN o no se incluyó transcriptasa reversa.

Luego de 40 ciclos de amplificación, los productos se visualizaron en geles de agarosa. El primer resultado correspondió a la amplificación con el cebador específico de G13 (figura 9 A.). No se logró la amplificación de un solo producto, sino dos bandas principales, sobre un fondo de productos de amplificación.

Un cálculo basado en el tamaño de las proteínas P estimado por western blots y comparación con proteínas P de otros organismos (8, 9, 10, 17), sugería que la banda de 420 pb podría codificar una proteína P completa (aproximadamente 110 aminoácidos) incluyendo las regiones no codificantes.

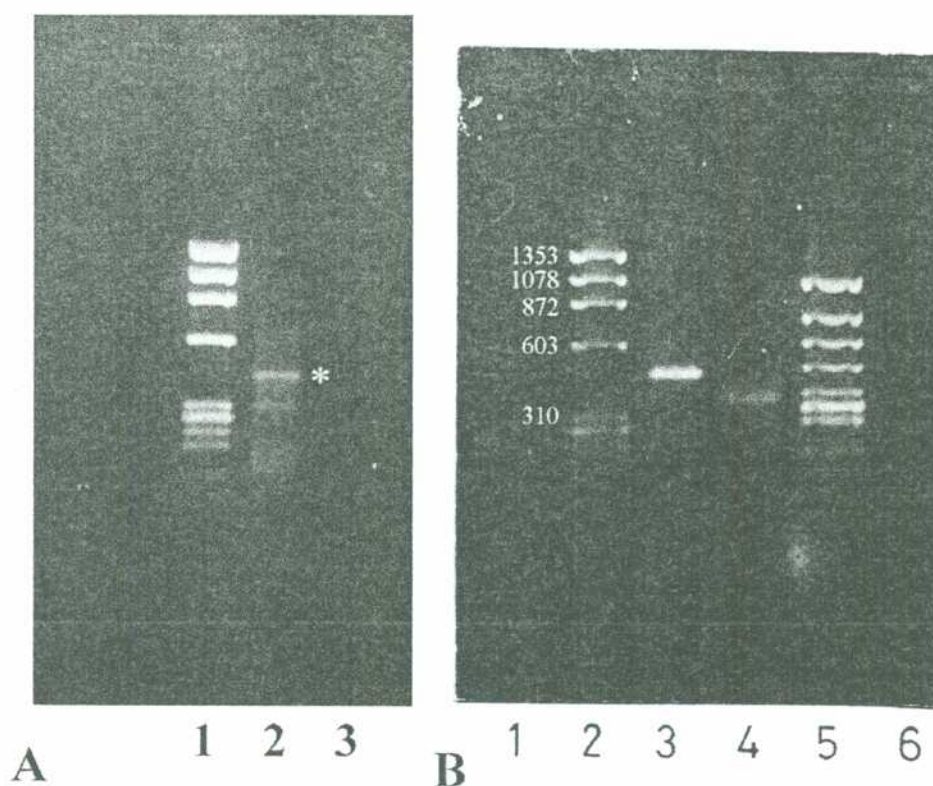
Para confirmarlo, se utilizó una técnica de PCR anidado (Nested PCR) (62).

Se diseñó un nuevo cebador ubicado a 50 bases "upstream" del cebador G13ext. Este se denominó G13int:

Figura 9-

A. Calle1, marcadores de PM; Calle2, Amplificación G13; Calle3, Reacción sin ARN. **B.** Calle1, Reacción sin ARN; Calles 2 y 5, Marcadores de PM; Calle3, Amplificación L6; Calle 4, Amplificación G13 Nested

5'- CCTTCTTGGCGTCAGCGGC- 3' T_m= 64°C



Así, si el producto correcto era la banda de 420 pb, entonces el producto de PCR debería ser de ~370 pb.

Simultáneamente, se realizó la amplificación para obtener la secuencia completa del ARNm Lyt6.

Los resultados se pueden ver en la parte B de la figura 9. La amplificación anidada indica que la banda de 420 pb corresponde al mensajero de la proteína G13 (calle 4). Para Lyt6 se obtuvo un único producto de ~490 pb (calle 3).

Estos dos fragmentos amplificados se clonaron en el vector M13mp10 con extremos romos como se indica en materiales y métodos. Los fagos recombinantes obtenidos se secuenciaron por la técnica de Sanger sobre las dos cadenas.

Efectivamente, los productos amplificados por ARN-PCR contenían las secuencias

Figura 10.
Secuencia total
de G13(MAT1)
completada con
la secuenciación
de la banda de
la calle 4 de
la figura 9. Se
indica
recuadrado el
SL, doble
subrayado los
codones de
iniciación y
Stop y
subrayado el

SL

```

AACGCTATTA TTGATACAGT TTCTGTACTA TATTGCTCGA AGTCTCTTTA CGCCACCTCA 60
GCCATGTCCA TGAAGTACCT CGCCGCATAC GCCCTTGCGT CGCTGAACAA ACCGACACCA 120
GGCGCCGCCG ATGTGGAGGC CATCTGCAAG GCCTGCGGTA TCGAAGTTGA GAGCGATGCA 180
CTCTCGTTTG TCATGGAATC CATTGCCGGC CGGAGCGTTG CCACTCTCGT GGCGGAGGGC 240
GCGGCGAAGA TGAGCGCTGT TGCCGTCTCC GCTGCTCCTG CTGCCGGTGA TGCGGCCGCT 300
CCTGCTGCTG CTGCTGGCGG TGCCGCCGCC CCTGCCGCCG CTGACGCCAA GAAGGAAGAA 360
GAGGAGGAGG ACGATGACAT GGGCTTTGGT CTGTTTGA CTGTTTGA CTGTTTGA 420
TGCCGGAGCA GAAAGGGACT GCGTCACTGC GCACTTGAAC AAGGAAAAAC ATACTGGCCG 480
CCGTATTTGG GGTGTGCATG ACTGGAAGGC ATCTTGCAGT TCGACAACCT TTTTGTGATT 540
CTGTTTTTTC TATTTATGAA TGAATTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA 597
  
```

que completan las secuencias parciales de G13(MAT1) y Lyt6 (figuras 10 y 11).

Las secuencias tienen los siguientes números de acceso en el EBI nucleotide sequence database :

X65025 para G13(MAT1)

X65065 para Lyt6

De acuerdo con las secuencias obtenidas (ver más adelante), las proteínas se denominaron:

G13(MAT1), **TcP2 α**

Lyt6, **TcP1**.

PJL5 (ver el análisis de 5D en el capítulo anterior) se denominó **TcP2 β** .

Figura 11.
Secuencia total
de L6
completada con
la secuencia de
la banda de la
calle 3 de la
figura 9. Se
indica
recuadrado el
SL y subrayados
los codones de
iniciación y
Stop.

SL

```

1  AACGCTATTA TTGATACAGT TTCTGTACTA TATTCTGTGA GTCTCAGCCG CCCCCCGCTA
61  ATTTACTCTC AAGGGAAATG TCCTCCAAAC AGCAGCTTGC CTGCACCTAC GCCGCCCTGA
121 TTCTTGCCGA TAGCGGCAAG ACGGATATGG ACAGCCTGTT GAAAGTGACA AAGGCCGCCG
181 GTGTTGACGT CAGCAAAGGG ATGGCCTCGG CGTTTGCCAG CATCCTCAAG AACGTTGACA
241 TCAACGACGT GCTCTCCAAA GTGAGCTTTG GTGGTGTGTC TCCTGCTGCC GGTGGTGCCA
301 CCGCTGCTCC TGCTGCTGCT GCTGCTGCCG CCGCCCCTGC CGCCGCCGCC GCAAAGAAGG
361 AAGAGGAAGA GGAAGACGAC GATATGGGCT TTGGTCTGTT TGACTAGGCG CAGAAGCCCC
421 TTTCTTGAAG GATTTACACA CAAGTGGCAT GGTTTGTGTT TTGTGTTTGT TTTTCTCAAA
481 TATTTATTCA CTCAACAAAA AATGAGTATA TATATAAAAA AAAAAAAAAA AAAAAAAAAA
  
```

2.2- Análisis de las secuencias de ARNm de proteínas P de *T.cruzi*. Ubicación filogenética y nomenclatura.

Proteína TcP2 α

El ARNm completo de TcP2 α tiene 597 bases y codifica para una proteína de 112 aminoácidos con un PM teórico de 10965.36 Da. Tiene 34 alaninas (Ala) y 18 residuos ácidos (Asp, Glu), resultando pI estimado de 4.85. La gran mayoría de las Ala se encuentran agrupadas en una región denominada “bisagra flexible” presente en todas las proteínas P (17). TcP2 α tiene homología de 50% con TcP2 β , la otra proteína de tipo P2 de *T.cruzi*, (figura 12.A, Tabla I) (63). El codón AUG de iniciación esta precedido por 63 bases no codificantes que incluyen 35 bases del SL. Un segundo codón AUG esta presente a solo dos codones “downstream”. Si se considera este codón como iniciador, la secuencia TcP2 α puede alinearse exacta-

Figura 12.

A.
Comparación de
las P2 de
T. cruzi (Tc) con
las de levadura
(Y)

B.
Comparación de
la P1 de *T. cruzi*
(Tc) con la P1
humana (H) y de
levadura (Y)

A

1

10

20

30

Dominio N-Terminal

TcP2α

TcP2β

YP2α

YP2β

M S M K Y L A A Y - A L A S L N K P T P G A A D V E A I C K A

- - * * * * * - * * V G * S G G * * S K S A * * * V L * *

- - * * * * * * L L * N A A G - N * * D * T K I K * * L E S

- - * * * * * * L L * V Q G G N A A * S * * * I K * V V E S

40

50

60

TcP2α

TcP2β

YP2α

YP2β

C G I E V E S D A L S F V M E S I A G R - S V A T L V A E G A

A * V P * D P S R V D A L F A E F * * K - D E D * V C T * * K

V * * * I * D E K V * S * L S A L E * K - * * D E * I T * * N

V * A * * D E A R I N E L L S * L E * K G * L E E I I * * * Q

70

80

90

Región Bisagra

TcP2α

TcP2β

YP2α

YP2β

A K M S A V A V S A A P A A G D A A A P A A A A - G G - A A

S * L V G G V T R P N A * T A S * P T A * * * * S S * - * *

E * L A * * P A - * G * * S A G G * * A * - - - S * D * *

K * F A T * P T - G G A S S A A * G * A G * * * - * * D * *

100

110

Dominio C-Terminal

TcP2α

TcP2β

YP2α

YP2β

A P A A A D A K K - - - E E E E E D D D M G F G L F D

* * * * * A - - - - - * * * * * * * * * * * * *

* - - - - E E E * E E E A A * * S * * * * * * * * *

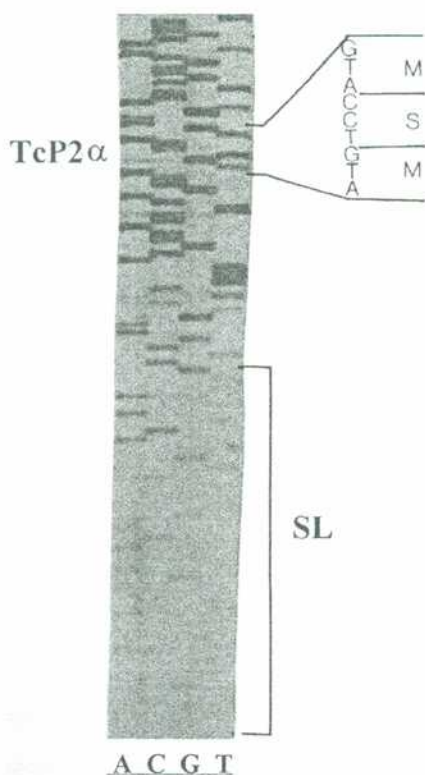
- - - - - E E E * E E E A K * * S * * * * * * * * *

B

| | | | | |
|------|--|---|---|----|
| | | Dominio N-terminal | | |
| | | 1 | 10 | 20 |
| TcP1 | | M - S S K Q Q L A C T Y A A L I L A D S G K T D M D S - L L | | |
| HP1 | | * A * V - S E * * * I * S * * Y * H * D E V * V E T D K I N | | |
| YP1α | | * - * T E S A * - - S * * * * * * * * * E T E I S S E K * * | | |
| | | 31 | 40 | 50 |
| TcP1 | | K V T K A A G V D V S K G M A S A F A S I L K N V D I N D V | | |
| HP1 | | A L I * * * * * N * G P F W P G L * * K A * A * * N * G S L | | |
| YP1α | | T L * N * * N * P D E N I W * D I * * K A * D G Q N L K * I | | |
| | | Region Bisagra | | |
| | | 61 | 70 | 80 |
| TcP1 | | L S K V | S F G G V A P A A G G A T A A P A A A A A A A P | |
| HP1 | | I C N * | G A * * P * * * * * A * P * G G P * P S T * * * * | |
| YP1α | | * V N T | S A - - - - - G * A A P A G V A G G V * G G E * G | |
| | | Dominio C-terminal | | |
| | | 91 | 100 | |
| TcP1 | | A A A A A | - - - K K E E - E E E D D D M G F G L F D | |
| HP1 | | * E E K K | V E A * * * * S * * S * * * * * * * * * | |
| YP1α | | E * E * E | K E E E E A K - * * S * * * * * * * * * * | |

mente con las proteínas P2. Sin embargo, el uso del primer AUG genera una extensión de dos residuos (Met-Ser) que es característico de secuencias N-

Figura 13.
Secuencia que
muestra la
extensión Met
(M)-Ser (S)
inusual de
TcP2 α . SL es la
secuencia líder
del ARNm



terminales de proteínas P1 de eucariontes inferiores (figura 13 y descripción de P1 más adelante).

Esta extensión Met-Ser no se encuentra en las P2 de otros eucariontes. Lamentablemente, el consenso de Kozak (64) no es aplicable a tripanosomas y los intentos de armar un consenso comparando varios genes de *T. cruzi*, no nos dio un resultado claro.

Efectivamente, con posterioridad a la publicación de nuestra secuencia, se conocieron las de *Leishmania infantum* y *Babesia Bovis* (65, 66). En estas secuencias también se puede observar la extensión de dos aminoácidos en el extremo N-terminal presente en TcP2 α . Estos hechos nos llevan a

postular que efectivamente el primer AUG es utilizado y esta extensión es única de parásitos protozoarios (figura 14).

Proteína TcP1

El ARNm de TcP1 tiene 541 bases y codifica para una proteína de 109 aminoácidos (67). Es la proteína P que concentra el mayor número de Ala en la región de "bisagra flexible", 16 sobre un total de 18 aminoácidos (figura 12 B.). TcP1 es una proteína ácida con un pI estimado de 5.13. Presenta un motivo aminoacídico interno GG(V/P)APAAG, posición de alineamiento 67-74, característico de proteínas P1 de eucariontes superiores, pero no de levadura o protozoarios (figura 12 B.).

Comparte 51% y 40% de homología con la P1 humana y la P1 de levadura, respectivamente y tiene la típica secuencia N-terminal de proteínas P1 (Tabla I).

Clasificación

Figura 14.
Comparación de
secuencias P de
distintos
organismos. Hs,
humana; Y,
levadura; Tc,
T. cruzi; Li,
Leishmania
infantum; Rn,
Rata; Tb,
T. brucei.

| Dominio Globular N-terminal | | | | | | Dominio C-terminal | | | | |
|--|----------------------------------|------------|------------|------------|----------|--|---|---|------------|--|
| | 1 | 10 | 20 | 30 | 40 | ... 95 | * | * | 115 | |
| HsP1 | MASVSELAC | IYSALILEDD | EVTVTEDKIN | ALIKAGVNV | EPF W... | EAKKEESEE | S | | DDDMGFGLFD | |
| Yp1 α | MST--ESAL | SYAALILADS | EIEISSEKLL | TLTNAANVPD | ENI W... | KEEEEAKEE | S | | DDDMGFGLFD | |
| Yp1 β | MS---DSDI | SFAAFILADA | GLEITSDNLL | TITKAAGANV | DNV W... | EEKEEEAEE | S | | DDDMGFGLFD | |
| TcP1 | MSSKQQLAC | TYAALILADS | C-KTDMDSL | KVTKAAGVDV | SKG N... | AAAKKEEEE | E | | DDDMGFGLFD | |
| | 1 | 10 | 20 | 30 | 40 | | | | | |
| HsP2 | --MRYVAS | LLAALGGNSS | PSAKDIKKIL | DSVGIEADDD | R | ... DEEKEESEE | S | | DDDMGFGLFD | |
| YP2 α | --MKYLAAY | LLLNAAGN-T | PDATKIKAIL | ESVGIEIEDE | K | ... EKEEEAEE | S | | DDDMGFGLFD | |
| TcP2 α | MSMKYLAAY | ALASLNE-PT | PGAADVEAIC | KACCIEVESD | A | ... ADAKKEEEE | E | | DDDMGFGLFD | |
| LiP' | MSTKYLAAY | ALASLSK-AS | PSQADVEAIC | KAVHIDVDQA | T | ... AKKDEPEEE | A | | DDDMGFGLFD | |
| BbL12 | MALKYVSSY | LLAVAAGNEN | PSVDDLKIL | DAVGSVDVEE | C | ... KKPEAEPEE | E | | EDDMGFSLFD | |
| YP2 β | --MKYLAAY | LLLQGGNAA | PSAADIKAVV | ESVGAEVDEA | R | ... AKEEEAKEE | s | | DDDMGFGLFD | |
| TcP2 β | --MKYLAAY | ALVCLSG-GT | PSKSAVEAVL | KAAGVPVDPS | R | ... PAAAAEEEE | E | | DDDMGFGLFD | |
| LIP | --MQYLAAY | ALVALSG-KT | PSKADVQAVL | KAAGVAVDAS | R | ... KKEEPEEEE | A | | DDDMGFGLFD | |
| Consenso eucariótico P EESDDDM-GFG---LFD | | | | | | | | | | |
| HsP1 | -AGGPAPAAGAAPAGGPAPSTAAAPAEK. | | | | | KVEAKKEESEEESDDDM-GFG---LFD | | | | |
| P2 | -ASVPAGGAVAVSAAPGSAAPAGSAPAAAEK. | | | | | KDE-KKEESEEESDDDM-GFG---LFD | | | | |
| P0 | -AAPAAAAAPA--- | | | | | KVEA-KEESEESEEDDM-GFG---LFD | | | | |
| RnP1 | -AGGPAPAAGAAPAGGPAPSAAAAAPAEK. | | | | | KVEAKKEESEEESDDDM-GFG---LFD | | | | |
| P2 | -ASVPAGGAVAVSAAPGSAAPAGSAPAAAEK. | | | | | KDE-KKEESEEESDDDM-GFG---LFD | | | | |
| P0 | -LAP-AAAAPA. | | | | | KVEAK-EESEESEEDDM-GFG---LFD | | | | |
| TcP1 | -FGGVAPAAGGATAAPAAAAA. | | | | | AAAPAAAA-AKKEEESEEDDM-GFG---LFD | | | | |
| P2 α | -GGVTRPNAATASAPTAASASSG. | | | | | AAAPAAAA-EEEEEDDM-GFG---LFD | | | | |
| P2 β | -SAVAVSAAPAGDAAAPAAAAGG. | | | | | AAAPAAA-DAKKEEESEEDDM-GFG---LFD | | | | |
| P0 | | | | | | -TGA. AAAPAAAAEP--EEEDDDDF-GMGA--LF- | | | | |
| TbP2 β | | | | | | -PSGGAAAHAAA.ASAPAAAAA--EAEEEDDDDM-GFG---LFD | | | | |

Las proteínas TcP1, TcP2 α (G13), TcP2 β (P-JL5) se compararon con las secuencias de rata, *Drosophila* y levadura (13, 16). El análisis de matrices las divide en dos grandes grupos P1 y P2 : TcP2 α (G13) se relaciona con YP2 α de levadura y P-JL5 con YP2 β , TcP1 se agrupa dentro de las proteínas grupo P1 (ver tabla I).

En concordancia, propusimos una clasificación de las proteínas P parasitarias basada en la nomenclatura de Wool et al. (68) (tabla 2). Por esta razón, como la proteína de G13 (MAT1) se ubica en el grupo de las proteínas P2, se la denominó TcP2 α . P-JL5 se denominó TcP2 β por agruparse con las P2. En esta clasificación la

Tabla I. Homología de proteínas P

| P1 | | | | | | P2 | | | | | |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | Rn | Dm | YPα | YPβ | TcP1 | Rn | Dm | YPβ | YPα | TcP2β | TcP2α |
| P1 | | | | | | | | | | | |
| Rn | 1.000 | 0.607 | 0.434 | 0.368 | 0.440 | 0.298 | 0.292 | 0.282 | 0.330 | 0.308 | 0.259 |
| Dm | | 1.000 | 0.509 | 0.396 | 0.413 | 0.330 | 0.330 | 0.291 | 0.349 | 0.308 | 0.259 |
| YPα | | | 1.000 | 0.538 | 0.387 | 0.255 | 0.255 | 0.226 | 0.274 | 0.245 | 0.236 |
| YPβ | | | | 1.000 | 0.415 | 0.283 | 0.311 | 0.311 | 0.255 | 0.292 | 0.274 |
| TcP1 | | | | | 1.000 | 0.330 | 0.312 | 0.257 | 0.264 | 0.346 | 0.321 |
| P2 | | | | | | | | | | | |
| Rn | | | | | | 1.000 | 0.611 | 0.564 | 0.547 | 0.383 | 0.375 |
| Dm | | | | | | | 1.000 | 0.491 | 0.472 | 0.411 | 0.366 |
| YPβ | | | | | | | | 1.000 | 0.641 | 0.402 | 0.391 |

primera o dos primeras letras indican el organismo de procedencia, Tc, *Trypanosoma cruzi*; Li, *Leishmania infantum*, etc (tabla 2).

Es notorio que las tres proteínas de bajo peso molecular de *T.cruzi* conservan la secuencia C-terminal sin serinas (Ser) EEEEDDDMGFGLFD (figura 5), desviándose del consenso C-terminal P. La típica Ser del C-terminal es reemplazada por un residuo glutámico (Glu), cambio no conservativo que afecta la antigenicidad de estas proteínas (29, 30, 31). El análisis filogenético reveló que las proteínas P de *T.cruzi* se encuentran entre las mas divergentes de todos los organismos, ubicán-

dolas en una de las bifurcaciones mas profundas del dendrograma (figura 15). De hecho la separación del grupo P2 de *T.cruzi* en P2 y P2 parece ser una de las mas antiguas: las proteínas P2 y P2 en *T.cruzi* solo tienen un 50% de homología entre ellas, mientras que en levadura alcanzan el 64% (tabla I). Las proteínas P2 de *Tetrahymena* aparecen mas cerca de la raíz del árbol a causa de la no conservación de la

Tabla 2. Nomenclatura unificada propuesta para las proteínas ribosomales P de parásitos.

| Parasitos protozoarios | Nomenclatura propuesta | Sinónimos |
|------------------------|------------------------|-----------|
| <i>T.cruzi</i> | TcP1 | TcP1, P2 |
| <i>T.cruzi</i> | TcP2a | TcP2b |
| <i>T.cruzi</i> | TcP2b | PJL5, |
| <i>T.cruzi</i> | TcP0 | TcP0 |
| <i>T.brucei</i> | TbP2b | TbP2 (a) |
| <i>L.infantum</i> | LiP2a | LiP' |
| <i>L.infantum</i> | LiP2b | LiP |
| <i>bovis B.</i> | BbP2a | BbL121A |
| <i>B.bovis</i> | BbP0 | BbL10 |

(a) S. Kunz y N. Muller, pers. comun.

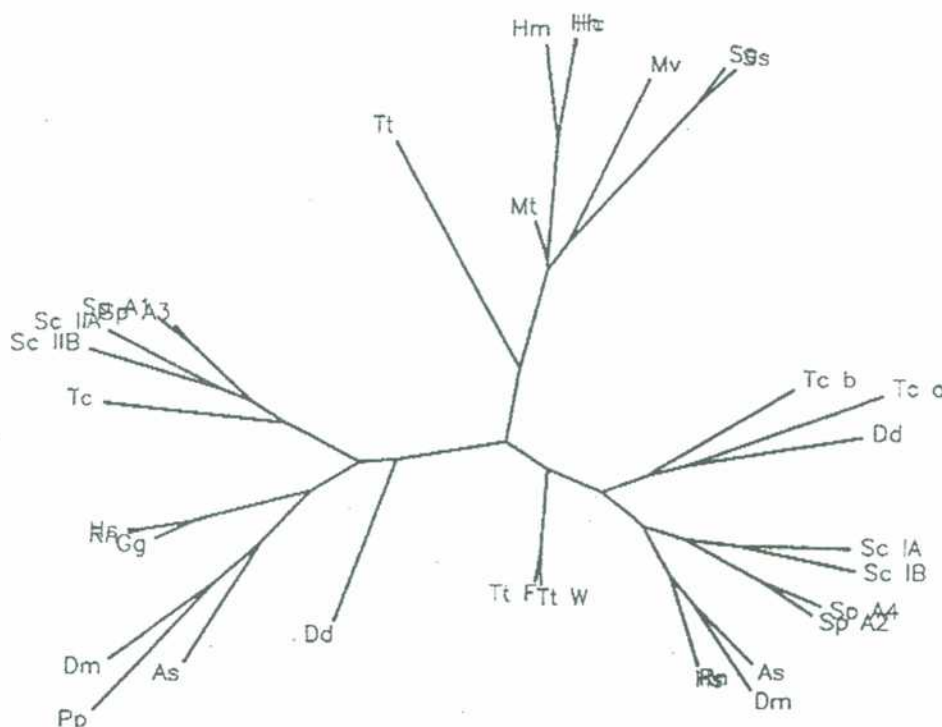
zona C-terminal (20) (figura 14), mientras que las proteínas P de bajo peso molecular de *T. cruzi* mantienen la conservación del extremo C-terminal consenso.

Alineando mas de 30 proteínas P de diversos organismos se observa que:

- Todas las P2 (en posición de alineamiento 40, figura 14) conservan un aminoácido básico, Lys o Arg.
- Las P1 conservan en la misma posición un residuo Trip.

Sin embargo, las P2 de parásitos protozoarios no tienen el residuo básico en

Figura 15-
Dendrograma de proteínas P construido con el programa PROTPARS de la serie PHYLIP. A la izquierda se ubican las P1 y a la derecha las P2. La única P de bajo PM de archaeobacterias se ubica hacia arriba.
Tc=TcP1,
Tcb=TcP2 α
Tca=TcP2 β ,
Tt=Tetrahymena termophila,
Dd=Dictiostelium discoideum,
Sc=levadura,
Dm=Drosophila melanogaster,
As=Aetmia salina.



posición 40, mientras que TcP1 y TtP1, no conservan el Trip.

Otro residuo conservado en general es la Leu en posición 48. Tampoco en este caso TcP2 α , TcP2 β , TtP2 α y TtP2 β lo mantienen.



ANALISIS DE LA EXPRESION DE LAS PROTEINAS P.

3.1- Análisis de las proteínas P por transcripción y traducción *in vitro* de los ARNm clonados.

Objetivos:
El objetivo de este capítulo es analizar las secuencias P clonadas por traducción in vitro y estudiar la expresión de sus ARNm

El hecho de que existan en eucariontes inferiores dos proteínas P2 (α y β) y dos proteínas P1 (α y β) sugería que posiblemente *T. cruzi* tendría un sistema similar. Sin embargo, pese a aislar dos proteínas P2, solo encontramos una P1. Estos datos podían indicar que no existe otra proteína P1 o bien que es tan divergente que no reacciona con los anticuerpos y las sondas correspondientes (el ejemplo es TcP0). Sin embargo, los datos de western blot parecen indicar la existencia de mas proteínas P que las caracterizadas.

Se decidió entonces transcribir y traducir IN VITRO los ARNm de las proteínas TcP2 α , TcP2 β y TcP1 para constatar si sus productos correspondían a las bandas detectadas en western blots de ribosomas, dado que modificaciones post-traduccionales y la conformación de estas proteínas podrían afectar su movilidad electroforética (7).

Transcripción y traducción IN VITRO

Para realizar la transcripción IN VITRO se clonaron los ADNc de las proteínas TcP2 α , TcP2 β y TcP1 en el plásmido pGEM4Z (Promega, USA) bajo dirección de promotor como SP6 o T7. Así se construyeron los plásmidos pTcP1SP6, pTcP1T7 y p2 β SP6, p2 β T7. Para clonar TcP2 α se diseñó un nuevo cebador 5'

próximo al ATG de inicio, y se realizó una amplificación sobre ADN genómico usando como cebador 3' G13ext. Esto requirió poner a punto las condiciones de PCR dada la inespecificidad que se obtuvo con las primeras condiciones ensayadas (ver detalles en materiales y métodos y figura 16). Cuando se logró amplificar una sola banda, esta se clonó en pGEM4Z (p2αSP6 y p2αT7).

Los plásmidos se usaron para sintetizar los ARNm correspondientes, tal como se describe en materiales y métodos, pag..

Los ARN así sintetizados se tradujeron IN VITRO en con un sistema de reticulocitos de conejo (Amersham, USA). La inclusión de Met marcada en el ensayo permitió analizar los productos de la síntesis por autorradiografía. El proceso se esquematiza en la figura 17 y los detalles en materiales y métodos, pag...

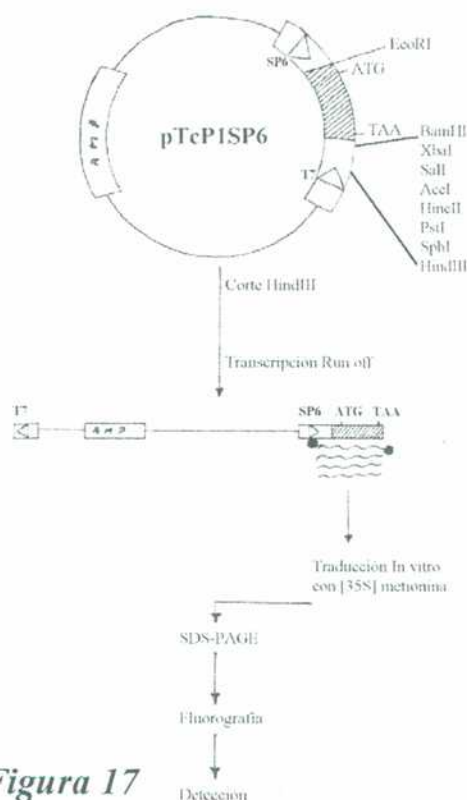


Figura 17

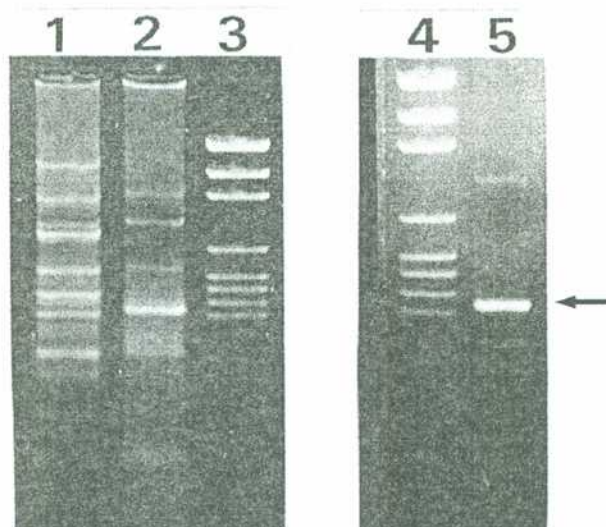


Figura 16-
Amplificación de la
región codificante com-
pleta de TcP2α 1,
Reacción sin adi-
cionales. Tm:50°C. 2,
Reacción con 10%
DMSO, 50°C. 3 y 4,
Marcadores de PM. 5,
igual a 3, pero a 60°C.
La flecha indica la
banda esperada de 370
pb

Las tres proteínas P analizadas (TcP1, TcP2α y TcP2β) migraron de acuerdo con lo esperado. La migración de los productos neosintetizados fue comparable con la migración de las proteínas P del ribosoma analizadas por western blot, (PM estimado entre 10 y 12 Kd) (Figura 18). Se confirmó que este resultado representa la proteína completa sintetizada in vitro por dos criterios: A) amplificando a partir de los ARN sintéticos toda la región codificante y B) verificando

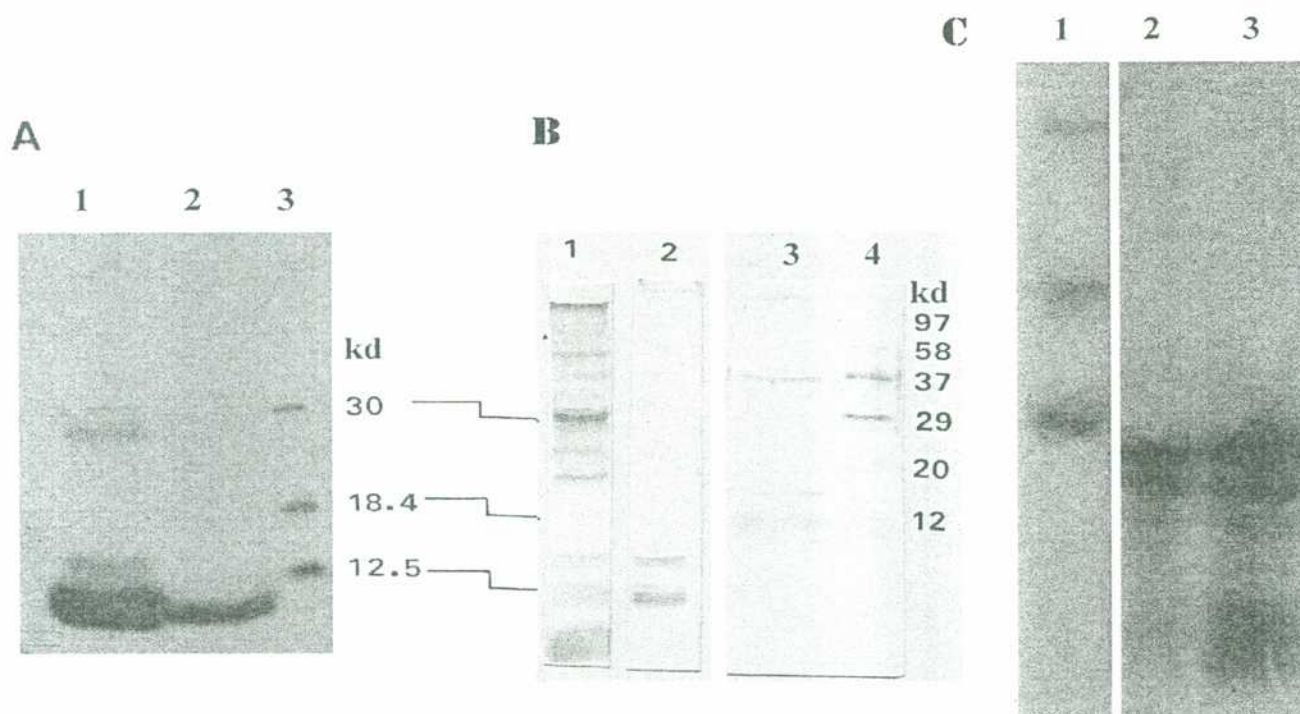


Figura 18-

A. 1, Producto TcP1 traducido in vitro. 2, Immunoprecipitación de 1 con Ac. de suero lúpico P+. 3, marcadores de PM [14 C] (Amersham, USA).

B. Western blot de ribosomas de *T. cruzi*. 1, se usó suero de un paciente chagásico crónico. 2 y 3, se usó suero de pacientes lúpicos P+. 4, marcadores de PM.

C. 1, marcadores de PM [14 C] (Amersham, USA). 2 y 3, Productos TcP2 β y TcP2 α respectivamente traducidos in vitro.

por inmunoprecipitación con anticuerpos anti-P de paciente lúpico la proteína neosintetizada (precipitación positiva implica conservación del extremo C-terminal de la proteína) (figura 18 A., calle 2). Así parece ser que la proteína de 17 Kd que aparece en western blots no ha sido clonada todavía (figura 18 B., calle 3).

Los western blot revelados por suero lúpico P+ (figura 18 B., calles 1, 2 y 3) muestran 3 bandas en la zona de bajo peso molecular (alrededor de 12 Kd), además de una proteína de 17 Kd y la P0 de 37 Kd. En principio estas tres bandas corresponden a las tres proteínas clonadas, y queda abierta la pregunta sobre la identidad de la banda de 17 Kd.

La reacción de esta proteína con los anticuerpos antiP define su pertenencia a la familia P ribosomal, pero su secuencia aminoacídica no es conocida y en todas las experiencias de clonado no existe ninguna secuencia P que corresponda a este tamaño.

3.2- Análisis de la expresión de ARNm por clonado usando ARN-PCR y por clonado usando una biblioteca de expresión en λ gt11

Con el objetivo de clonar la proteína P de 17 kd y continuar el análisis de la expresión de las otras se emplearon dos estrategias:

A) Amplificación por ARN-PCR de todas las posibles proteínas ribosomales de bajo peso molecular usando como cebador 3' una secuencia consenso de la región 3' que codifica para el extremo C-terminal, esperando poder amplificar los ADNc de todas las proteínas P presentes en *T.cruzi* (con excepción de P0).

B) Rastreo sobre una biblioteca de expresión de ADNc de *T.cruzi* estadio tripomastigote sanguíneo (P. Levy-Yeyati, tesis doctoral, UBA, 1992), usando suero de pacientes lúpicos con alto título anti-P. Se esperaba identificar todas las proteínas que contienen un motivo P ribosomal en la región C-terminal.

Usando la **estrategia A** se amplificaron de 3 bandas de 350, 400 y 750 pb (figura 19). Esta imagen se asemeja a la observada en western blots, dado que las bandas de 350 y 400 pb de ADNc amplificadas podrían codificar para proteínas de alrededor de 12 Kd (observar la calle 2 de la figura 19, en donde se amplificó el ADNc de TcP1 como control). La banda de 750 pb, podría generar una proteína de alrededor de 20 Kd.

Los productos se clonaron en el fago M13 y se seleccionaron 21 fagos recombinantes para su secuenciación.

Las secuencias demostraron que las bandas de 350 y 400 pb estaban

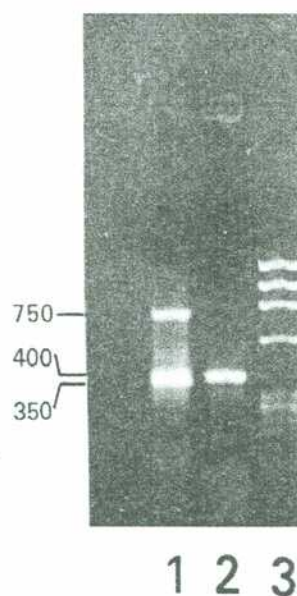


Figura 19-
1, ARN-PCR sobre ARN total de *T.cruzi*, usando un cebador consenso 3' P y el cebador del SL. 2, Amplificación de TcP1 con los mismos cebadores como control. 3, marcadores de PM. Se indican los tamaños de las 3 bandas amplificadas.

Figura 20.

A. Secuencias de amplificación inespecíficas que corresponden a regiones 5' de ARNm. Subrayado en negro se señalan los sitios de clonado SmaI. Subrayado en marrón, la secuencia del SL. Recuadrados se muestran los cebadores usados en la reacción de PCR. El clon C es un producto de amplificación inespecífico debido a ambos cebadores.

B. Nueva variante de TcP2 β obtenida en la misma reacción de PCR que los productos de A. Se recuadra en azul el sitio HindIII mutado.

mayoritariamente representadas por los ADNc de TcP2 β (6 secuencias) y TcP1 (5 secuencias) respectivamente. Otras seis secuencias eran productos inespecíficos de amplificación y cuatro eran delecciones. Estas secuencias, figura 20 A., corresponden a regiones 5' de ARNm desconocidos identificados como tales por presencia de SL y el ATG de inicio.

A.**Clon A**

CCCCAACGCTATTATTGATACAGTTTCTGTACTATATTGTGAGAGAGGGAGGCAAGAAGGTGGATTGCGTAGGTAGGCGT
 AAGTAATAGACGTGCACGTCGAAACTGGAAATAGTTTTTTTTCTTTCTTTGCGA ATG GCG AGC AAA CAG ACG
 CCA CAC ACC GAT GGC GAG AAC TCT TCA GCG GGG AAA ATT AGA GGA CTT GAG M A S K Q T
 P H T D G E N S S A G K I R G L E
 ATGACATGGGATTGGTCTTTTTGACGGGG

Clon B

CCCCAACGCTATTATTGATACAGTTTCTGTACTATATTGACCACAAGCTAACTTAAGGATTTTTCTCAACGTTACAC
 ATG ACG CGT CAA GCA GCA ATT TCC GAG GACATGGGATTGGTCTTTTTGACGGGG
 M T R Q A A I S E

Clon C

CCCCAACGCTATTATTGATACAGTTGTGTTTCAATCCAGTTTATTGGGTGATCCATTATCTTGCGTTCTAATTTATGATAT
 GTGTTTTGTTTTGGTGGTATCTATATTATGTGGTTTTTTTGAATTGTATTAGTAAGTGTATAATACATTCTGGTTGTTT
 GACATGGGATTGGTCTTTTTGACGGGG

B.**TcP2 β -RA**

CCCCAACGCTATTATTGATACAGTTTCTGTACTATATTGTTGTCTTGTGTCAG ATG AAG TAC CTT GCC GCT TAC
 GCT CTT GTT GGC CTC TCA GGG GGG ACC CCC TCA AAG TCT GCA GTT GAG GCC GTA TTG AAG M K Y L A A Y
 A L V G L S G G T P S K S A V E A V L K
 GCT GCC GGT GTC CCT GTC GAT CCT TCT CGT GTG GAT GCA CTT TTT ACC GAG TTT GCG GGC
 A A G V P V D P S R V D A L F T E F A G
 AAG GAC TTT GAT ACG GTA TGC ACT GAA GGC AAA TCG AAA CTT GTA GGA GGG GTT ACA CGT
 K D F D T V C T E G K S K L V G G V T R
 CCC AAT GCT GCC ACC GCT TCT GCA CCC ACT GCC GCT GCC GCC GCC TCT AGT GGT GCC GCC
 P N A A T A S A P T A A A A A S S G A A
 GCG CCT GCT GCA GCT GCG GAA GAG GAG GAG GAC GAT GAC ATG GGA TTT GGT CTT TTT GAC
 A P A A A A E E E E D D D M G F G L F D
 GGGG

No se obtuvo ninguna secuencia de TcP2 α . Para TcP2 β se obtuvo la secuencia de una variante, figura 20 B., que no habia sido descripta previamente (A. Schijman, tesis doctoral, UBA, 1993). Esta variante presenta 6 sustituciones nucleotídicas con respecto a las ya clonadas que determinan un polimorfismo en un sitio interno HindIII y un cambio aminiacídico en posición 43 (figuras 7 y 20 B). No se obtuvo ninguna secuencia correspondiente a la banda de 750 pb. También fracasó el intento de clonar la banda. Finalmente, la secuencia directa sobre el templado PCR (banda de 750 pb) dio una lectura parcial de la region 5' que no mostró homología con ninguna proteína P. Sin embargo, los datos no descartan que esta banda pueda codificar para la proteína P de 17 kd.

Un resultado importante de esta experiencia de ARN-PCR se relaciona con el nivel de expresión de los genes P. Teniendo en cuenta que los transcritos de TcP1, TcP2 α y TcP2 β tienen la misma probabilidad de amplificarse por poseer:

longitudes aproximadamente similares

homología de 100% entre cebadores y secuencias de los ARNm,

se puede concluir que la cantidad de secuencias amplificadas podría estar relacionada con el nivel de expresión de los ARNm correspondientes. De acuerdo con esto, TcP2 α tendría el menor nivel de expresión de las tres proteínas P.

Usando la **estrategia B**, se rastreó una biblioteca de ADNc en λ GT11 del estadio tripomastigote sanguíneo de *T. cruzi*. Estos tripomastigotes se obtuvieron de ratones infectados con alta parasitemia (P. Levy-Yeyati tesis doctoral, UBA, 1992). El rastreo se realizó con un suero lúpico con altos niveles de anticuerpos anti-P. Se aislaron 12 fagos recombinantes denominados RAL. Con los fagos purificados se realizaron ensayos de placas de lisis con suero de pacientes chagásicos con alto título de anticuerpos anti-P. Este análisis permitió clasificar los fagos en tres grupos:

TABLA III

| | ANTICUERPOS ANTI-P+ | ANTICUERPOS ANTI-P + | SONDAS DE ADN | SONDAS DE ADN |
|--------------|---------------------|----------------------|----------------------------|---------------|
| FAGOS (pb) | SUERO CHAGASICO | SUERO LUPICO (SLE) | TcP2 α , TcP0, TcP1 | TcP2 β |
| RAL1 (220) | + | ++++ | - | - |
| RAL2 (450) | ++++ | ++++ | - | + |
| RAL3 (450)* | ++++ | ++++ | - | + |
| RAL5 (220) | + | ++++ | - | - |
| RAL6 (320)* | ++++ | ++++ | - | + |
| RAL8 (220) | + | ++++ | - | - |
| RAL10 (380)* | + | ++++ | - | - |
| RAL11 (300)* | + | ++++ | - | - |
| RAL12 (450) | ++++ | ++++ | - | + |
| RAL13 (350)* | ++++ | ++++ | - | + |
| RAL14 (420) | - | ++ | - | - |
| RAL15 (320) | ++++ | ++++ | - | + |

* fagos recombinantes cuyos insertos fueron secuenciados.

A) 6 fagos recombinantes que reaccionaron intensamente con ambos sueros.

B) 5 fagos recombinantes que reaccionaron débilmente con suero chagásico, pero intensamente con suero lúpico.

C) 1 fago recombinante que no reaccionó con el suero chagásico, y lo hizo débilmente con el lúpico.

Cuando el mismo ensayo se hizo usando sondas radiactivas de los ADNc de proteínas P, solo TcP2 β reconoció todos los fagos ubicados en el grupo A (tabla III).

Fagos de los grupos A y B se subclonaron y secuenciaron (marcados en la tabla III), el grupo C no se considero de importancia para el análisis en

esta tesis. Los 3 fagos secuenciados del grupo A corresponden a distintas variantes de TcP2 β .

Sorprendentemente, las dos secuencias obtenidas del grupo B codificaban para porciones de proteínas ribosomales P1 de ratón (98 y 34 aminoácidos respectivamente, figura 21). Esto se explica porque los tripomastigotes se obtuvieron de sangre de ratones infectados. En estas preparaciones de parásitos, siempre existe un cierto grado de contaminación con células sanguíneas del huésped (en particular glóbulos blancos), razón por la cual hay ARNm de ratón representados en la biblioteca.

*** DNA TRANSLATION ***

| | | |
|-----|---|-----|
| 1 | GAA TTC CAC GAC GAG GTG ACG GTC ACG GAG GAT AAG ATC AAT GCC | 48 |
| 1 | E F H D D E V T V T E D K I N A | 16 |
| 49 | CTC ATT AAA GCA GCT GGT GTC AGC GTC GAA CCT TTC TGG CCT GGC TTG | 96 |
| 17 | L I K A A G V S V E P F W P G L | 32 |
| 97 | TTT GCC AAG GCT CTG GTC AAT GTC AAC ATT GGG AGC CTC ATC TGC AAT | 144 |
| 33 | F A K A L V N V N I G S L I C N | 48 |
| 145 | GTA GGG GCT GGT GGG CCC GCT CCA GCA GCT GGA GCT GCG CCT GCT GGT | 192 |
| 49 | V G A G G P A P A A G A A P A G | 64 |
| 193 | GGT GCT GCT CCA TCC ACT GCC GCC GCC CCA GCT GAG GAG AAG AAA GTG | 240 |
| 65 | G A A P S T A A A P A E E K K V | 80 |
| 241 | GAG GCA AAG AAG GAA GAG TCC GAG GAG TCT GAA GAT GAC ATG GGC TTC | 288 |
| 81 | E A K K E E S E E S E D D M G F | 96 |
| 289 | GGT CTT TTT GAC TAA ACT GCT TTT GTT AAG TTG GCT AAT AAA GAG CTG | 336 |
| 97 | G L F D | 100 |
| 337 | AAC CTG TAA AAA AAA AAA AAA AAA AAA AAA AAA AAG AAT TC | 380 |

Esta proteína P de ratón, desconocida hasta la fecha, tiene todas las características de las proteínas P de bajo peso molecular de mamíferos. Su región C-terminal es 100% homóloga a la humana. La reactividad diferencial de estos recombinantes es de importancia para el análisis de los fenómenos autoinmunes en la enfermedad de Chagas. En efecto, los sueros de pacientes chagásicos reaccionan débilmente con las proteínas P de ratón pero muy intensamente con las protei-

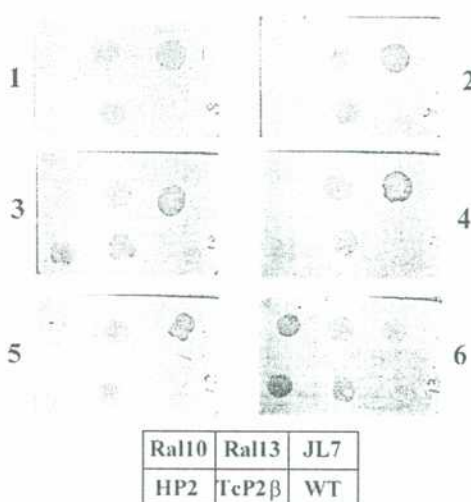


Figura 21-

Secuencia del ADNc de la proteína P1 de ratón aislada con el suero lúpico. Subrayada en negro se muestra la secuencia consenso P. Subrayada en verde está la señal consenso de poliadenilación eucariota. Recuadrados se muestran los sitios EcoRI de clonado.

Figura 22-

Paneles de fagos recombinantes probados con distintos sueros. 1-5 sueros de pacientes chagásicos con cardiopatía. 6. suero lúpico P+. JL7, antígeno flagelar que reacciona con todos los sueros chagásicos crónicos. HP2, P2 humana. WT, fago λ gt11 salvaje.

nas P del parásito, como lo demuestran los recombinantes del grupo A. Teniendo en cuenta que los anticuerpos anti-P de pacientes chagásicos y los autoanticuerpos anti-P de pacientes lúpicos tienen afinidad por los 13 aminoácidos C-terminales de las proteínas P del parásito y del hospedador (ver introducción), la débil reacción del suero chagásico con la proteína P mamífera (figura 22) solo puede explicarse comparando las secuencias C-terminales de las proteínas correspondientes:

P1 ratón: EE *S* EDDMGFMLFD

TcP2 β : EE *E* DDDMGFGLFD

La comparación revela que la única gran diferencia entre ambas es el cambio no conservativo de una serina por un ácido glutámico. Experiencias inmunológicas utilizando las proteínas recombinantes y péptidos sintéticos confirman que, tal como se observó en el rastreo inmunológico, la secuencia C-terminal de la proteína mamífera es menos antigénica frente al suero chagásico que la secuencia C-terminal de la proteína del parásito (D.Kaplan, tesis doctoral, UBA, 1995). De estos resultados se desprende que si bien existe un cierto nivel de autoinmunidad anti-P en sueros chagásicos, esta auto reactividad es diferente de la observada en sueros lúpicos. En Chagas entonces la reactividad auto anti-P sería producto de una reacción que, en un primer término, esta dirigida contra la proteína ribosomal P del parásito (31).

Si bien de este rastreo no se obtuvo ningún fago que codificara para la proteína de 17 Kd, las conclusiones obtenidas fueron igualmente importantes.

El conjunto de las experiencias A y B permite conclusiones adicionales. Parece existir una diferencia en la expresión de los genes de las proteínas P. Obsérvese que en A se trabajo con ARN de epimastigotes y no se obtuvieron ADNc de TcP2 α y en la experiencia B trabajando con tripomastigotes sanguíneos no se obtuvieron ADNc de TcP2 α y TcP1.

3.3- Análisis de la expresión de ARNm por Northern blot usando distintas condiciones de temperatura en parásitos *IN VIVO*.

Para investigar diferencias en la expresión de las proteínas P, se realizaron ensayos de Northern blot y Dot blot. La fuente de ARN en estos ensayos fueron epimastigotes que se sometieron a un tratamiento de shock térmico, uno de los factores necesarios para imitar el pasaje del epimastigote, que vive a 28°C en el intestino de la vinchuca, a la sangre del hospedador mamífero, a 37°C. Los epimastigotes se crecieron a 28°C hasta alcanzar la fase exponencial tardía en medio líquido LIT, y luego se procedió como se muestra en el siguiente esquema.

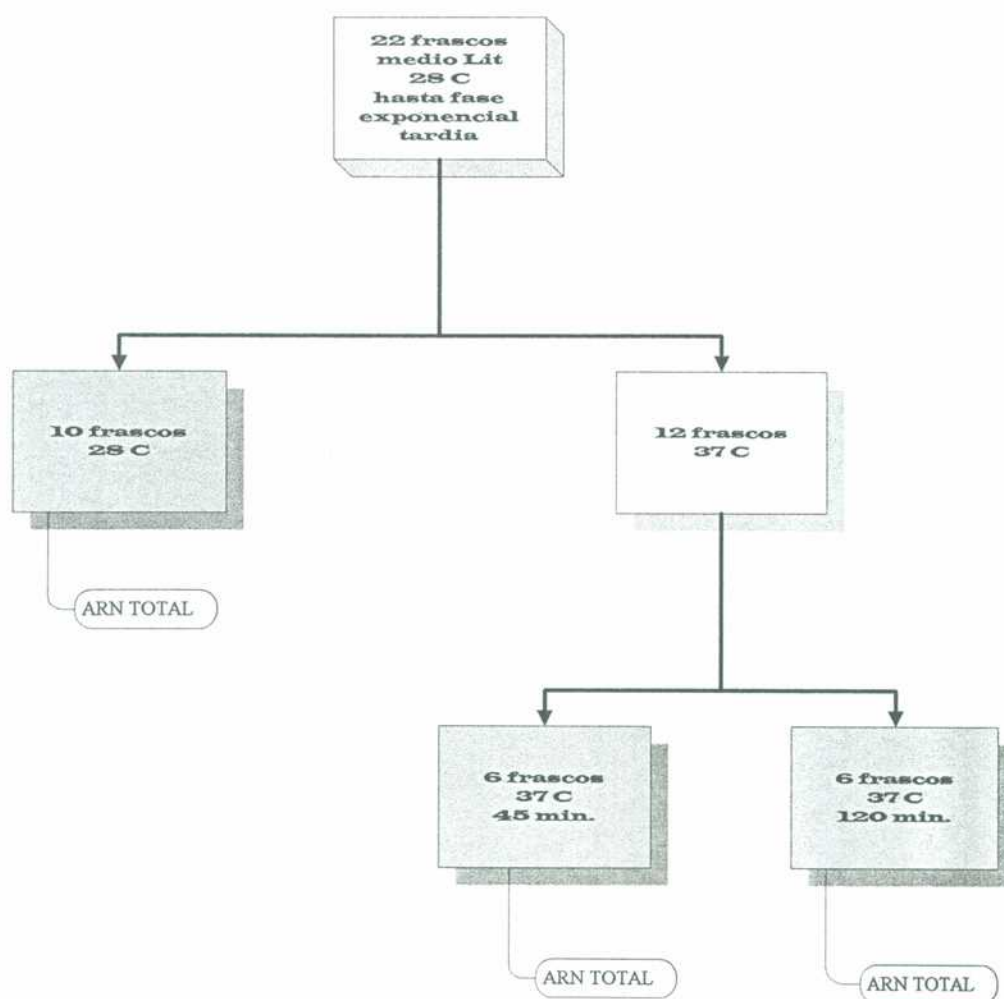


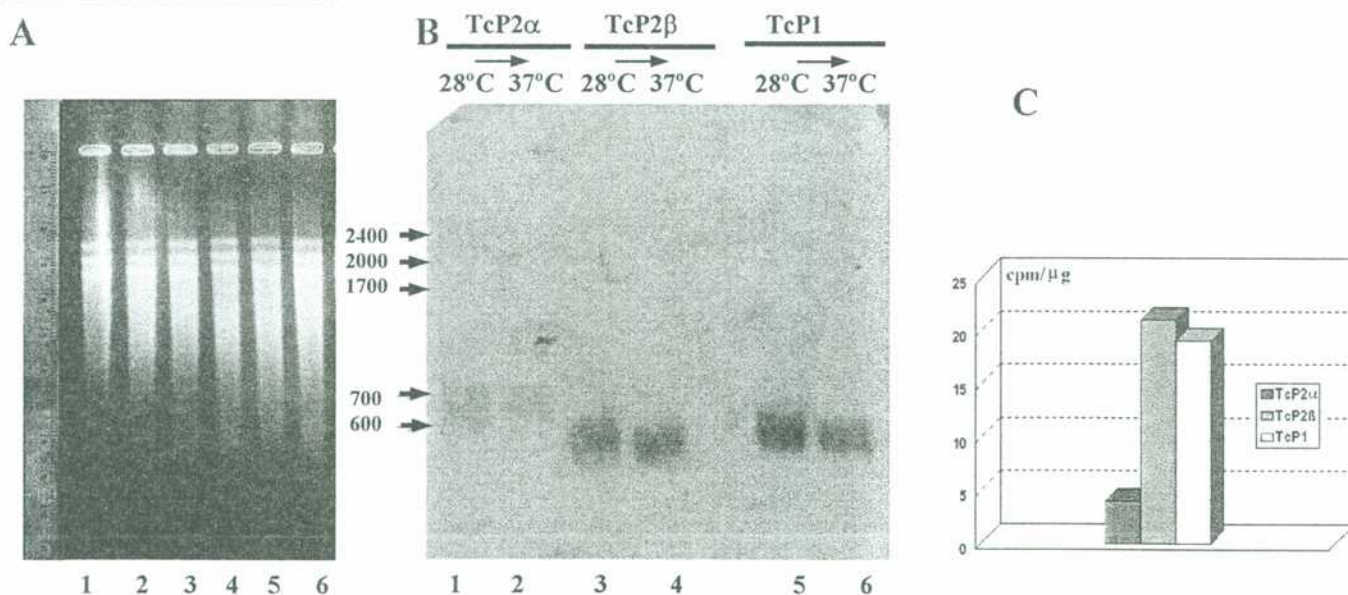
Figura 23-

A. ARN total de *T. cruzi* extraído de distintas condiciones de cultivo.

B. Gel de **A**, transferido e hibridizado con TcP2 α , TcP2 β y TcP1, respectivamente. Calles 1, 3 y 5, parásitos cultivados a 28°C; calles 2, 4 y 6, parásitos que recibieron el golpe de calor a 37°C por 2 horas. **C.** Gráfico de los resultados de Dot blot. Se grafica la pendiente (cpm/ μ g) de cada una de las rectas que resultan para cada sonda.

Luego de las incubaciones, los parásitos se cosecharon rápidamente, se resuspendieron en GTC y se preparo ARN total de cada condición de cultivo. Los ensayos de cuantificación por Dot blot se realizaron sembrando cantidades crecientes de ARN total. El ARN sintético de TcP1 se uso como control de concentración. Las sondas utilizadas fueron TcP1, TcP2 α y TcP2 β para las proteínas P. Las sondas usadas como control fueron la subunidad menor del ARN ribosomal de *T. cruzi* como control de expresión constitutiva (obtenida como se explica en el capítulo 1, a partir del fago 5D, banda B1) y la sonda HSP70 (P.Levy-Yeyati, tesis doctoral, UBA, 1992) como control de inducción por calor. El resultado se obtuvo por visualización directa luego de autorradiografía y por conteo de cpm en centelleo líquido. Como complemento de la cuantificación se realizaron ensayos de Northern blot, sembrando cantidades exactamente iguales de ARN total en cada calle.

Los resultados obtenidos indican que, en epimastigotes cultivados a 28°C, la expresión de los genes TcP1 y TcP2 β alcanzan niveles similares. Mientras que el nivel de expresión de TcP2 α es, aproximadamente, 5 veces menor al de los otros genes (figura 23).



Cuando se comparan los datos con los controles de cuantificación (hibridación en dot blot con cantidades conocidas de ARN sintético de TcP1), se puede estimar un porcentaje relativo de abundancia de los ARNm de cada proteína. Si consideramos que el ARNm es un 5% del ARN total, resulta lo siguiente:

***Abundancia relativa de ARNm de proteínas ribosomales
P en el ARN total de epimastigotes de T.cruzi***

TcP1 = 0.10 %

TcP2 β = 0.15%

TcP2 α = 0.02%

Se puede observar que TcP2 α presenta niveles de ARNm que se encuentran un orden por debajo de los niveles de los otros dos mensajeros (figura 23, A y B). Este hecho podría explicar por que no se aislaba el ADNc de TcP2 α en las experiencias descriptas en 3.2.

En relación a los niveles de ARNm en los parásitos incubados a 37°C, podemos observar en la figura 23 A., que no hay diferencias importantes entre las dos temperaturas, excepto por una pequeña disminución en el nivel de mensajeros de TcP1 a 37°C. Es un resultado esperable dado que se trata de proteínas que son expresadas constitutivamente.

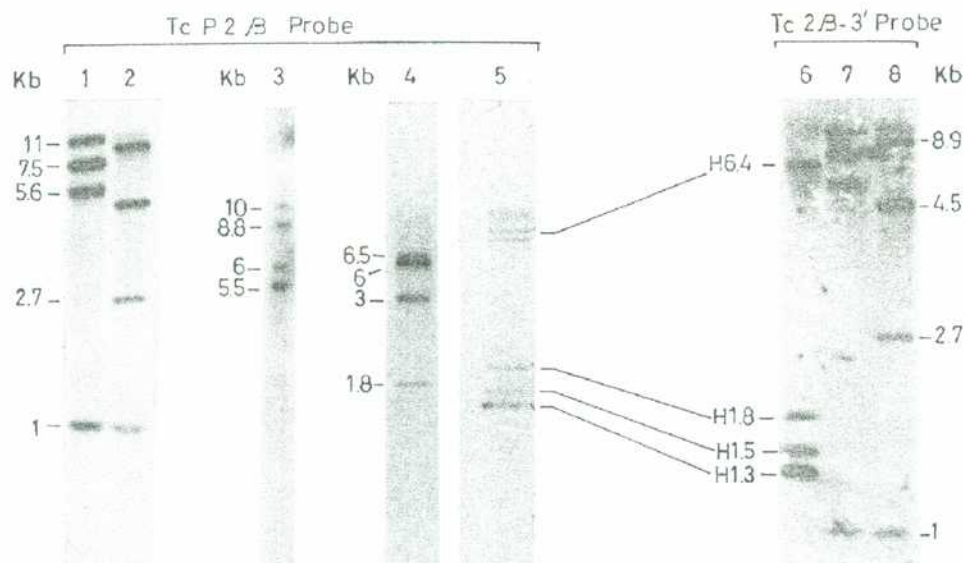
Cuando estos resultados se analizan conjuntamente con los datos de Southern blot se obtiene un panorama aun más interesante. En efecto, la proteína TcP2 β se encuentra codificada por al menos 4 genes por genoma haploide en la cepa Tulahuen de T.cruzi (ver figura 24, esto será analizado en capítulos posteriores). La proteína TcP2 α es codificada por 2 copias génicas, como se puede observar al mapear por southern blot el ADN genómico usando cortes dobles. Esto se confirma tanto con enzimas que cortan en sitios fuera del gen y cuya localización desconocemos, como cuando se utiliza la enzima HinfI como punto común de anclaje que corta en la zona 3' no codificante, en todos los casos se obtienen dos bandas (figura 25, las flechas señalan las dos copias identificadas). El mapeo realizado por

Figura 24-Southern Blot TcP2 β

1 y 7, EcoRI. 2 y 8,

BamHI. 3, KpnI. 4,

Sall. 5 y 6, Hind III.

**Figura 25-**Southern blot TcP2 α

1, BamHI. 2, BamHI +

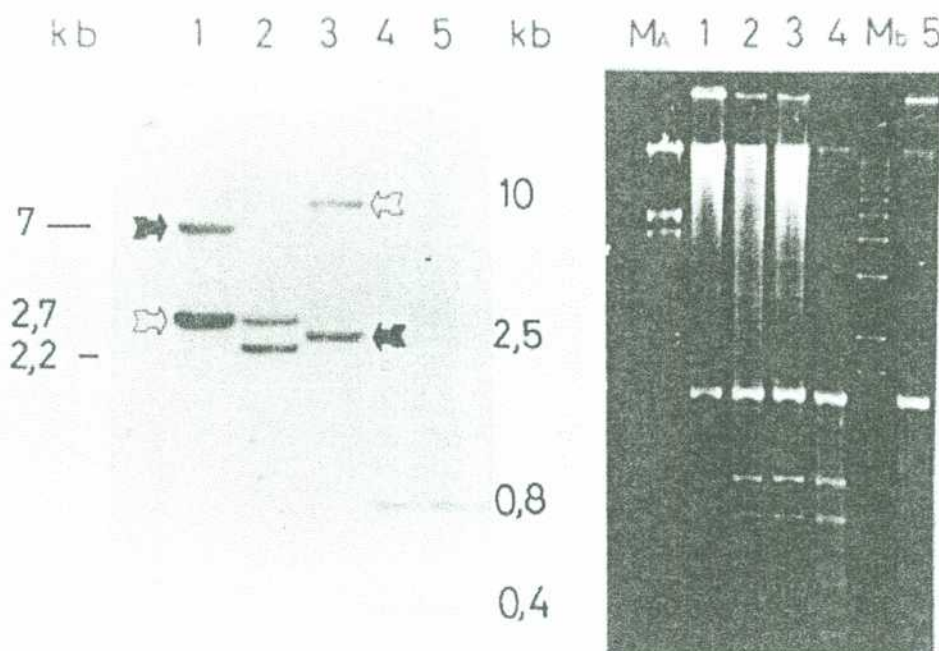
EcoRI. 3, EcoRI. 4,

HinfI. 5, HinfI +

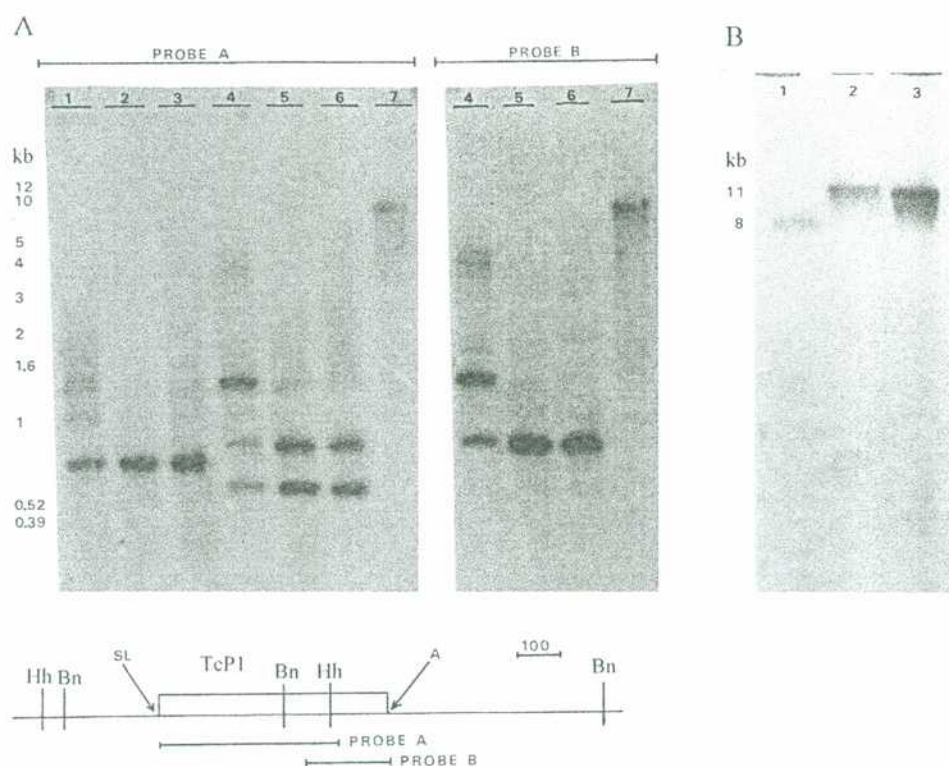
EcoRI.

Las flechas blancas y
negras indican bandas
relacionadas entre si.

Se muestra tambien el
gel de agarosa en
donde se ven los cortes
dobles



southern blot para la proteína TcP1 indica que se encuentra codificada por un solo gen. En la figura 26, se observan los detalles del mapeo y las sondas utilizadas. El corte con enzimas de reconocimiento de seis bases (EcoRI, BamHI, HindIII y Sall) mostró en todos los casos una única banda. Para descartar la posibilidad de que esta banda contuviera un tandem de copias del gen TcP1, se realizó una experiencia de mapeo con cortes parciales

**Figura 26-**

Southern blot TcP1.

A. 1, 2 y 3, digerido con HhaI por 10, 30 minutos y 4 horas respectivamente. 4, 5 y 6, digerido con BnI por los mismos tiempos que HhaI. 7, digerido por Sall.

B. 1, Hind III. 2, EcoRI. 3, BamHI.

El diseño de las sondas se muestra en el esquema inferior. Hh, HhaI; Bn, BnI; SL, spliced leader; A, adición de poliA.

usando enzimas de corte frecuente (ver detalles en la leyenda de la figura). Se utilizaron dos sondas (A y B, figura 26); la sonda A es cortada internamente por BanI. En un corte total de ADN genómico con BanI hibridizado con sonda A, se esperaría encontrar al menos tres bandas si hubiera mas de una copia en tandem contenida en ese locus; en cambio, se esperarían solo dos bandas si solo existiera una copia. De la misma forma, la hibidación con la sonda B sobre cortes parciales de BanI debería identificar múltiplos de bandas si existiera un tandem, mientras que la distribución de bandas sería azarosa en caso de estar presente una sola copia. Observando la figura 26, resulta claro que existe una sola copia del gen TcP1 contenida en el locus definido por las enzimas EcoRI, BamHI, HindIII y Sall. El mismo resultado se confirma usando la sonda A en un corte parcial con la enzima HhaI (figura 26, A, calles 1, 2 y 3).

De estos resultados se concluye que TcP2 β y TcP1 tienen niveles de expresión de ARNm similares a pesar de estar codificados por distinto numero de genes. TcP1 con un único gen logra los mismos niveles de

expresión que TcP2 β codificada por al menos 4 genes (por el clonado de expresión sabemos que todas las copias se transcriben) (69, 70). Mientras que TcP2 α codificada por al menos dos genes mantiene un nivel de expresión muy bajo. Es probable que las señales regulatorias de estos genes estén influenciando su expresión. Se sabe que los tripanosomas ejercen un fino control de la regulación de genes a través del mecanismo de Transplicing y no a nivel de promotores (ver introducción) (43, 44). Es probable que existan diferencias a nivel de esas señales y que estén modulando la regulación de la expresión.

CAPITULO 4

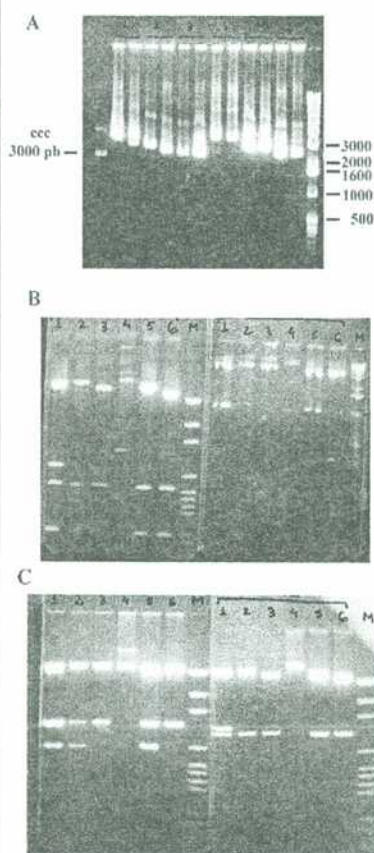
Objetivos:
el objetivo de este capítulo es analizar en detalle la organización genómica de los genes TcP2 β

Figura 27-
A. Escisiones de los plásmidos contenidos en los fagos λ ZAP TcP2 β +. **B y C.** Mapeo de restricción de los plásmidos 1 a 6 con EcoRI, BamHI, HindIII, PstI y la combinación de esas enzimas de restricción

CLONADO DE GENES DE PROTEINAS RIBOSOMALES P.

4.1- Organización genómica de proteínas P

Teniendo en cuenta los resultados descriptos en el capítulo anterior, se decidió profundizar el estudio de la organización genética de los genes de las proteínas P. El objetivo de este capítulo fue caracterizar y entender las pautas que rigen su expresión. La idea que inspiró este enfoque partió de la observación que TcP1, con un solo gen, logra los mismos niveles de expresión que TcP2 β , expresado a partir por lo menos 4 genes y casi un orden mas que TcP2 α transcripto por 2.



Clonado de los genes TcP2 β .

Se comenzó con el clonado genómico de los genes TcP2 β . Estos se clonaron de una biblioteca genómica de T.cruzi cepa Tulahuen construida en el vector λ ZAPII.

El rastreo se realizó utilizando una sonda de ADN completo de TcP2 β (320 pb). Luego del rastrear 150.000 fagos, se aislaron 10 recombinantes positivos, denominados λ G1 a 10, para su posterior análisis.

Los fagos recombinantes fueron purificados y el ADN clonado fue escindido IN VIVO aprovechando las particularidades del vector λ ZAPII (71). Los

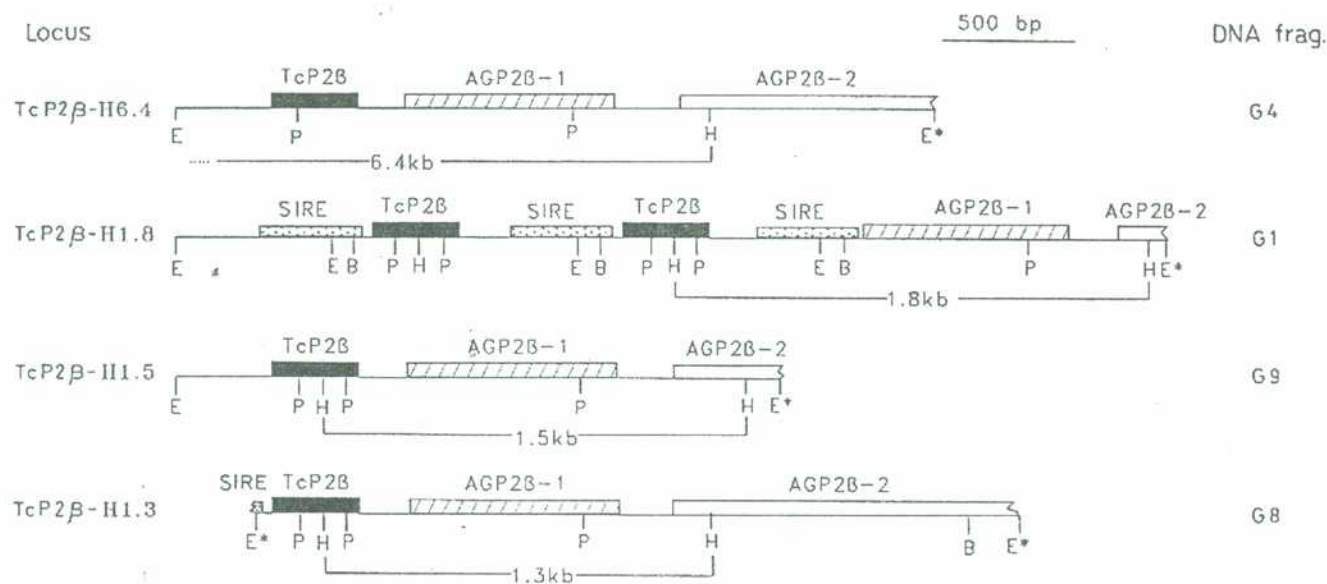
plásmidos recombinantes escindidos fueron preparados y analizados por corte con una batería de enzimas de restricción.

Las enzimas fueron elegidas de modo que no cortaran más de una vez dentro de la secuencia plásmidica, y con un sitio de corte en la zona codificante de TcP2 β (información obtenida a partir del ADNc). Así, el ADN del plásmido se digirió con las siguientes enzimas: HindIII, EcoRI, PstI, KpnI, SacI, BamHI, y combinaciones dobles de las mismas. En la figura 27 se muestra el mapeo de algunos de los plásmidos.

En base a las digestiones que se realizaron, se fueron deduciendo los mapas de restricción de cada uno de los fagos genómicos. Este análisis indicó que los 10 fagos recombinantes aislados se podían agrupar en cuatro grupos según los fragmentos obtenidos por restricción. Este resultado concordaba con los datos obtenidos por análisis de Southern Blot de ADN genómico de *T. cruzi* cepa Tulahuen II, figura 24 en capítulo 3. En este

Figura 28-

Mapa de los 4 loci genómicos que contienen a TcP2 β E, EcoRI. B, BamHI. P, PstI. H, HindIII. E*, EcoRI de clonado. Los puntos indican que el fragmento 6.4 kb excede los límites de clonado. Se indican los tamaños de las bandas HindIII que le dan el nombre a cada uno de los loci



análisis se puede observar que el corte con varias enzimas de restricción da resultados coincidentes con la existencia de 4 loci genómicos diferentes para TcP2 β . Estos loci fueron denominados de acuerdo con el tamaño de las bandas Hind III del Southern Blot: HI1.3, HI1.5, HI1.8, HI6.4. Estas bandas Hind III fueron usadas para determinar la identidad de cada uno de los locus

clonados. Así, de los fagos seleccionados, λ G1 con un fragmento de 3.8 Kb contenía al locus **H1.8**; λ G4 de 2.8 Kb contenía al locus **H6.4**; λ G8 de 2.7 Kb contenía al locus **H1.3**; y λ G9 de 2.3 Kb al locus **H1.5**.

Estos fagos recombinantes genómicos fueron secuenciados en su totalidad.

La combinación de los datos de secuencia y de mapeo con las enzimas de restricción, permitió deducir el mapa físico para cada uno de los loci que contienen al gen TcP2 β y su entorno genómico (figura 28). La secuencia correspondiente a cada uno de ellos se muestran en las figuras 29, 30, 31, 32. En la figura 33, se muestra un análisis comparativo de las secuencias

Estas secuencias nucleotídicas de los genes TcP2 β codificada en cada locus mostró que es una familia de genes altamente homólogos, pero no idénticos. Las diferencias entre ellos son consistentes con mutaciones puntuales, la mayoría transiciones, que generan en algunos casos variaciones menores en la secuencia proteica (70). Los transcritos generados desde los loci H1.5 y H1.3 se corresponden con los descriptos anteriormente (A. Schijman, tesis doctoral, UBA, 1993). El transcripto generado a partir de H1.8 es el descripto en el capítulo 1 secuenciado desde la banda B2 del fago 5D y el transcripto de H6.4 corresponde al ARNm de TcP2 β descripto en el capítulo 3, este es el gen que presenta polimorfismo en el sitio interno HindIII.

En los fragmentos clonados, Los genes TcP2 β están ligados a dos genes de función desconocida.

Los genes TcP2 β están ligados a otros dos genes en cada uno de los locus estudiados. De 5' a 3', el ordenamiento de los genes es el siguiente: TcP2 β , región intergenica de 190 pb, **AGP2 β -1** (Associated P2 β gene-1), región intergenica de 180 pb, **AGP2 β -2**. Rio arriba de los codones de iniciación de AGP2 β -1 y AGP2 β -2 se encuentra una clásica secuencia polipirimidinica de 20 y 24 bases respectivamente seguidas de los correspondientes

```

1 ---GAATTC CTCTCTAGGA GAATTTTTTG AGCACAACGA TTTAGAGAAT AGTTTAGCTG
61 AAAGGAGTGA TGCTTGCTTG GATGAGTGCA AGACCATGGG TGACGTAAAT GCCCTATTCT
121 CTGAAAGCGG ATTGCTTACC TTGCCCTTAA CACGAGAGGC GGTTAACGAG TTTCTTCAAA
181 ATATGCACGG GATTCACTAA GAAACATTTT CACGCACGAA AGCGAAATTA TTATGATTGT
241 TATTATAATA CTTTTTCTTT GTTGTTTTAT CCACTTATTA TTGTGTGTGT AAATTGTTTT
301 TACCTTTTTT CTTTCCAAC TTCTTTTATG ATGTCTTTTC TTTTTTTTTT TTTGCTCTA
361 TAAGTTGTCT TGTCAGATGA AGTACCTTGC CGCTTACGCT CTTGTTGGCC TCTCAGGGGG
421 GACCCCTCA AAGTCTGCAG TTGAGGCCGT ATTGAAGGCT GCCGGTGTCC CTGTCGATCC
481 TTCTCGTGTG GATGCACTTT TTACCGAGTT TGCGGGCAAG GACTTTGATA CGGTATGCAC
541 TGAAGGCAAA TCGAAACTTG TAGGAGGGGT TACACGTCCC AATGCTGCCA CCCTTCTGCT
601 ACCCACTGCC GCTGCCGCCG CCTCTAGTGG TGCCGCCGCG CCTGCTGCAA CTGCGGAAGA
661 GGAGGAGGAC GATGACATGG GATTTGGTCT TTTTGACTAA AGACAAATGC TTATTTTATT
721 TTATTTTGA TTTTGTCTCT CTGTTGTCCA TGAAGAGGCG CTGGAGTCCT GTGAATAAAA
781 TGCTAGAGAG TAAATAAAAG TGGGAACAGT AACTGCTGCA CTCAAACGCA TGACAAAAAA
841 GGGTACGCCA TTGGGTGGC CTTTTTTTTT TTTTCCCTTT TAAGTTTTGT TCTATGACC
901 CTTGAGACAC GTAGAGACAA GTACGATACA AGCCAGCTGA CGCTCTCGTC GTTAGTTGCA
961 CTTTGCGCGG ATTGTTTCAG CTTTTGTTTG GAAAGGACCG CGAGAGTTCT GGACAATCAA
1021 AGCCGCGCAC CAAAAGGATG GGAAATCAG TTTGCGCGG TTGCCAAACA ACGTCGACAA
1081 GAGGACTGCG AACTTTGTGA ACGTTACTGC CATCGAGTTT CATCTGTTTG CCCTGCCTCG
1141 AAGGATATCG AACTTTTGA GGCTTTGGGA TATAAGGAGG TAGCAAAAAG GACTTCTCTA
1201 GATGGATTAC CCACGGTTCA GCAGATGTCT AGGTGCGAGA ACGCTGTTTC GGAGTTCTCA
1261 AGAGTTCTTT TTTCACCTGC GCACAAGCCT AGTAAGGTGG CATTGAAGA TGTTGACCCG
1321 AAAAGTGGCT TGAAAAAAG TGAGGAAGGT GGTGATGGTG AGCCAACCGA GCTTCATAAA
1381 AAGGTACTTA GTGGTGGATC AATAGATCAC GCCATCCTGC AGCGTTGGAG CACTGCTAGG
1441 AAGGGACGAG CCGAAATGGA GGAGATGGGG GTGCTTCGAA GTGCGGAAGA GACCGCATCG
1501 CTCAGGCAGG AACTTGGACG CCAGAGGTCC GAAGAGTACA TTAAGACAAG GGGGACGACT
1561 CCGAAGTGGT GAGCAAGCAC GAAGAAGCAC AAACAAGCA AAACACAACA AGAACAACA
1621 GTTTTATTTT TATTTTTTTT TTCTCTCTCT CTCTTTTTTT TGGTCTTCTC TTCTATATAT
1681 TGGTAGCAAC TCTTTTTTCC TGCCTCTCTT TCTATATATA ATTTTTTATT TGTTAGTTTT
1741 TCTTTACCAT GCAGAGTAAG GTTACGCGAG ATGTGCGCAT AGTTTTTTTT TTATCGATCG
1801 ATATGGTAGC CGGATGTGTG GCGGGAAAGA CTTATTACT TCACGTAAAG CGGGGAGACA
1861 AGATTCACGC GTCGTCGCCT CCGCTCATTG CAGACGATGG AGGAAAGCTT CACGTGAAC
1921 TGTCTCGCAT CTTTACGTCA ACGCTGCAAC GACAGAACGA TTGGAAGTAC AAGAAGAAGC
1981 CACTCAAGGT TCGCATTGAA GAGTTGCGCT CCCATTTTGT GACAAGTTTG TCCTACGACT
2041 TGAGCAAGAA CATTGTTGAA GAACTTTTC CGGAGAGAAG GGTGATTATC CGTGAACGGG
2101 ACGGGTTAAG AGAGGTTTCA TTAAGACTGA GGGGTGTAAG AGAAAGTCTT TATATCCGGC
2161 AACAAGATGT TCAAACTGCT TCCAATGCTT CCTCTAGCTC TTACATGCGC GACGCGAGCA
2221 CCGCGTCCTT GTTTGGACCC CACGATCCCT CTGAGGAACA TATTAATCAT TCAACCACTT
2281 CATTTTTGGT TGACGATTAC TTAAGTGTCA TTGATAGTTT GACAGGGACT GCAAGCGAGC
2341 AAAGGAGCGA GGATATCCAT TCTGAGACGA AAGAACTCA GGTGGACAAA GAACTCAATA
2401 AAATAAATTT TGACAAGTGT GCTTCAGCCC CGGATACATC TTTACACCCG TTAGTTGCAC
2461 CCATGCATCG AAGGACGGGG GCGTGCCAA ATACAATCAG CAAGGAAAAC TTTAAGAATC
2521 TTGCTTTTAC ACGGGTCTCA GATGCGGCTA TTTGAAGGG TAAGGATTTG TCCACTCTTA
2581 AGCCTTCTTG GAAGTGAAT ATTGCTTTTG AGGACCCCGT TCTAGGGGTC ATTAGGCAGA
2641 TTTGTGATGC TGGAAGGTGG TGTGGAGACA TTGGTAGGGA GTCTCAAACA GCGAGAGATG
2701 CTTGGAAGGG AGCTCATTTA GCTCTTGATT TTTATTTAAA TCAGGATGGA GAAAACCTGG
2761 GTCGCTTTGT TGTTAGTTTT CTGGCTCATG TTGTTGTTGA AGTGTGTCGA GAGCC

```

Figura 29-
*Secuencia del frag-
 mento genómico G4,
 que contiene al loci
 TcP2β-H6.4. Re-
 cuadrado en negro,
 TcP2β. Recuadrado en
 marrón, AGP2β-1. Re-
 cuadrado en verde,
 AGP2β-2. Subrayado
 se indican las regiones
 polipirimidínicas y
 además en negrita el
 aceptor AG*

Figura 30-
Secuencia del frag-
mento genómico G1,
que contiene al locus
TcP2β-H1.8.
En negrita se indica la
secuencia insertada.
Dentro de esta secuen-
cia, en *itálica* se mues-
tra la señal de trans-
splicing. En *negrita* y
subrayado, se indican
los codones de ini-
ciación y terminación
de los genes.

| | | | | | | | |
|------|-------------|-------------|-------------|-------------|------------|-------------|-------|
| 1 | GGAAATCCAC | TCTAGGAGAA | TTTTTTGAGC | ACAACGATTT | AGACAATAGT | TTTGCTGAAA | |
| 61 | GGAGTGATGC | TTGCTTGGAT | GAGTGCAAGA | CCATAGGTGA | CGTAAATGCC | CTATTCTCTG | |
| 121 | AAAGCGGATT | GCCTACCTTG | CCCCTAACAC | GAGAGGCGGT | TAAAGAGTTT | CTTCAAAAATA | |
| 181 | TGCAGCGGAT | TCACTAAGAA | ACATTTTCAC | GCACGAAAGC | GAAATTATTA | TGATTGTTAT | |
| 241 | TATAATACTT | TCTCTTTGTT | GTTTTATTCA | CTTATTATTG | TTGTATGAAA | TTGTTTTTAC | |
| 301 | CTTTTTTCTT | TTCCAACCTC | TTTTATGATG | TCTTTTCTTT | TTTTTTCTTA | GGAGCTTGTG | |
| 361 | AAAAGAATGA | TCGTGGGAGA | GCTGGCTAAC | TTAATTAATG | TATGTGTATA | TCCTGATAAA | SIRE |
| 421 | TGAATGCATT | CTTTATGATA | CTTCTACCG | TATGAATCTT | TTGGGAAGAA | CGCGACTTTG | |
| 481 | TAGGGGCGGG | AGCCGATAGA | GGCCAGATAA | TATTTTTAAT | TTTATTTTGC | CATCCCACCC | |
| 541 | ACCCCTTGA | TTCCCACCAC | GCGGCGGGGT | CTGTGGTTG | GAGGACCCCA | AAGTCTGCCA | |
| 601 | CTTCGTAAGT | AATAATATTT | CAAATCCCAA | CTGAGGACAA | AGGACCATGC | TAATGGTCCA | |
| 661 | CAGAATCTTA | TATATTGAAT | GAGAATAAAA | CATTGAGGCA | ATTAACATAA | GAACCTTTTT | |
| 721 | TCCTAACCCTG | TTTTGGTTTCC | ATAGATAATT | TCAGGATCCG | GCCAGCTGCC | CGGAGGATTA | TcP2β |
| 781 | TTCTGCTTTA | TAAGTTGTCT | TGTCAGATGA | AGTACCTTGC | CGCTTATGCT | CTTGTGGGCC | |
| 841 | TCTCAGGGGG | GACCCCTCA | AAGTCTGCAG | TTGAGGCGGT | ATTGAAGGCT | GCCGGCGTTC | |
| 901 | CTGTGATCCT | TTCTCGCGTG | GATGCACTTT | TTGCCGAGTT | TGCGGAGTAC | GACCTTGATA | |
| 961 | CGGTATGCAC | TGAAGGCAAA | TCGAAGCTTG | TAGGAGGGGC | TACACGTCCC | AATGCTGCCA | |
| 1021 | CCGCTTCTGC | ACCCACTGCC | GCTGCCGCGC | CCTCTAGTGG | TGCCGCCGCG | CCTGCTGCAG | |
| 1081 | CTGCGGAAGA | GGAGGAGGAC | GATGACATGG | GATTTGGTCT | TTTTGACTAA | AGACAAATGC | |
| 1141 | TTATTTTATT | TTATTTTGA | TTTTGTTTCT | CGGTTTTTCC | TGAAGAGGCG | CCGAGTCCCT | |
| 1201 | GTGAATAAAA | TGCGAGAGAG | TAAATAAAG | TAGGATCAGT | AACTGCTGCA | CTCAAACGCA | |
| 1261 | TGGCAAAAAA | GGGTACGCCA | TTGGGTTGGC | CTTTTTTTTA | GGAGCTTGTG | AAAAGAATGA | SIRE |
| 1321 | TCACGGGAGA | GCTGGCTAAC | TTAATTAATG | TATGTGTATA | TCCTGATAAA | TGAATGCATT | |
| 1381 | CTTTATGATA | CTTTCTACCG | TATGAATCTT | TTGGGAAGAA | ACCCCTTTTT | CCTAATCTGT | |
| 1441 | AACCGATAGA | GGCCGATATA | TATTTTTAAT | TTTATTTTGC | CATCCCACCC | ACCCCTTGA | |
| 1501 | TTCCCACCAC | GCGGCGGGGT | CTGTGGTTG | GAGGACCCCA | AAGTCTGCCA | CTTCGTAAGT | |
| 1561 | AATAATTTTC | AAATCCTAAC | TGAGGACAAA | GGACCATGCT | AATGGTCCAC | AGAATCTAT | |
| 1621 | ATATTGAATG | AGAATAAAG | ATTGAGGCAA | TTAACTAAAG | AACCTTTTTT | CCTAATCTGT | |
| 1681 | TTGGTTTCCA | TAGATAATTT | CAGGATCCGG | CCAGCTGCCC | GGAGGATTAT | TCTGCTTTAT | |
| 1741 | AAGTTGTCTT | GTCAGATGAA | GTACCTTGCC | GCTTATGCTC | TTGTTGGCCT | CTCAGGGGGG | |
| 1801 | ACCCCTCAA | AGTCTGCAGT | TGAGGCCGTA | TTGAAGGCTG | CCGGCGTTCC | TGTCGATCCT | |
| 1861 | TCTCGCGTGG | ATGCACTTT | TGCCGAGTT | GCAGGCAAGG | ACCTTGATAC | GGTATGCAT | |
| 1921 | GAAGGCAAT | CGAAGCTTGT | AGGAGGGGCT | ACACGTCCCA | ATGCTGCCAC | CGCTTCTGCA | |
| 1981 | CCCACTGCCG | CTGCCGCCGC | CTCTAGTGGT | GCCGCCGCGC | CTGCTGCAGC | TGCGGAAGAG | |
| 2041 | GAGGAGGACG | ATGACATGGG | ATTTGGTCTT | TTTGACTAAA | GACAAATGCT | TATTTTATTT | |
| 2101 | TATTTTGTAT | TTTTTTCTC | TGTTTTCCAT | GAAAGGCGGC | CGGAGTCCGT | TGAATAAAAT | |
| 2161 | GCGAGAGAGT | AAATAAAGT | AGGATCAGTA | ACTGCTGCAC | TCAAACGCAT | GGCAAAAAAG | |
| 2221 | GGTACGCCAT | TGGGTTGGCC | TTTTTTTTAG | GAGCTTGTGA | AAAGAATGAT | CACGGGAGAG | |
| 2281 | CTGGCTAACT | TAATTAATGT | ATGTGTATAT | CCTGATAAAT | GAATGCATT | TTTATGATAC | |
| 2341 | TTTCTACCGT | ATGAATCTTT | TGGGAAGAAC | GCGACTTTGT | AGGGGCGGGA | ACCGATAGAG | |
| 2401 | GCCGGATAAT | ATTTTAAATT | TTATTTTGCC | ATCCCACCCA | CCCCCTTGAT | TCCCACCACG | |
| 2461 | CGGCGGGGTC | TTGTGGTTGG | AGGACCCCAA | AGTCTGCCAC | TTCGTAAGTA | ATAATATTTT | |
| 2521 | AAATCCTAAC | TGAGGACAAA | GGACCATGCT | AATGGTCCAC | AGAATCTAT | ATATTGAATG | |
| 2581 | AGAATAAAGC | ATTGAAGCAA | TTAACTAAAG | AACCTTTTTT | CCTAATCTGT | TTGGTTTCCA | |
| 2641 | TAGATAATTT | CAGGATCCGG | CCAGCTGCCC | GGAGATTAT | CTCTTTTTTT | TTCCCTTTTA | |
| 2701 | AGTTTGTCTC | TTATGACCTT | TCAGACACGT | AGAGACAAGT | ACGATACAAG | CCAGCTGACG | |
| 2761 | CTCTCGTCGT | TAGTTGCACT | GTGCGCGGAT | TGTTTCGAGCT | TTTTTTTGGA | AAGGACCGCG | |
| 2821 | AGAGTTCTGG | ACAATCAAAG | CCGCGCACCA | AAAGGATGGG | AAAATCAGTT | TGCGGCGGTT | |
| 2881 | GCCAAACAAC | GTCGACAAGA | GGACTGCGAA | CTTTGCGAAG | GTTACTGCCA | TCGAGTTTCA | |
| 2941 | TCTGTTTGCC | CTGCCTCGAA | GGATATCGAA | CTTTTGGAGG | CTTTGGGATA | TAAGGAGGTA | |
| 3001 | GCACGAAGGA | CTTCCTTAGA | TGGATTGCCC | ACGGTTTCAGC | AGATGCTTAG | GTGCGAGAAC | |
| 3061 | GCTGTTTCGG | AGTTCTCAAG | AGTTCTTTTT | TCACTCGCGC | ACAAGCCTAG | TAAGGTGGCA | |
| 3121 | TTTGAGGATG | TTGACCCGAA | AAGTGGCTTG | AAAAAAGTGG | AGGAAGGTGG | TGATGGTGAG | |
| 3181 | CCAACCGAGC | TTCATAAAAA | GGTACTTAGT | GGTGGATCAA | TAGATGATGC | CATCTTGCA | |
| 3241 | CGTTGGAGCA | CTGCTAGGAA | GGGACGAGCC | GAAATGGAGG | AGATGGGAGT | GCTTCGAAGT | |
| 3301 | GCGGAAGAGA | CCGCATCGCT | CAGGCAGGAA | CTGGACGCTC | AGAGGTCCGA | AGAGTACATT | |
| 3361 | AAGACAAGGG | GGACGACTCC | GAAGTGGTGA | GCAAGCACGA | AGAAGCACAA | ACAAAGCAAA | |
| 3421 | ACACAACAAG | AACAACAAGT | TTTATTTTTTA | TCTCTCTCTC | TCTCTCTCTT | TTTTTTTTTG | |
| 3481 | GTCTTCTCTT | CTATATTATT | GGTAACAAC | CTTTTTTCCC | CCCTCTCTTT | CTATATATAA | |
| 3541 | TTTTTTCATTT | GTTAGTTTTT | CTTTACCATG | CAGAGTAAGT | TTACGCGAGA | TGTGCGCATA | |
| 3601 | GTTTTTTTTT | TATCGATCGA | TATGGTAGCC | GGATGTGTGG | CGGGAAGAC | TTATTTACTT | |
| 3661 | CAGCTTAAAG | GGGAGACAAG | ATTCAACCGG | TCGCTGCCTC | CGCTCATTGC | AGACGATGGA | |
| 3721 | GGAAAGCTT | | | | | | |

1 ----GGAATT CCACTCTAGG AGAATTTTTT GAGCACAACG ATTTAGACAA TAGTTTTGCT
 61 GAAAGGAGTG ATGCTTGCTT GGATGAGTGC AAGACCATAG GTGACGTAAA TGCCCTATTTC
 121 TCTGAAAGCG GATTGCCTAC CTTGCCCTTA ACACGAGAGG CGGTTAACGA GTTTCTTCAA
 181 AATATGCAGC GGATTCACCTA AGAAACATTT TCACGCACGA AAGCGAAATT ATTATGATTG
 241 TTATTATATAT ACTTTCTCTT TGTGTGTTTA TTCACTTATT ATTGTTGTAT GAAATGTITT
 301 TTACCTTTTT TCTTTTCCAA CTTCTTTTAT GATGCTTTTT CTTTTTTTTT TTTTGTCTTT
 361 ATAAGTTGTC TTGTCAGATG AAGTACCTTG CCGCTTACGC TCTTGTGTCG CTCTCAGGAG
 421 GGACCCCTCT AAAGTCTGCA GTTGAGGCCG TATTGAAGGC TGCCGGCGTC CCTGTCGATC
 481 CTCTCTCGCT GGATGCACCT TTTGCCGAGT TTTCAGGCAA GGACTTTGAT ACGGTATGCA
 541 CTGAAGGCAA ATCGAAGCTT GTAGGAGGGG TTACACGTCC CAATGCTGCC ACCGCTTCTG
 601 CACCCACTGC CGCTGCCGCC GCCTCTAGTG GTGCCGCCGC GCCTGCTGCA GCTGCGGAAG
 661 AGGAGGAGGA CGATGACATG GGATTTGGTC TTTTGTACTA AAGACAAATG CTTATTTTAT
 721 TTTATTTTTG ATTTTGTTC TGTGTTGTCC ATGAAGAGGC GCTAGAGTCC TGTGAATAAA
 781 ATGCGAGAGA GTAAATAAAA GTAGGAACAG TAACGTGTGC ACTCAAACGC ATGGCAAAAA
 841 AGGGTACGCC ATTGGGTTGG CTTTTTTTTT TTCCCCCTT TCAAGTTTGG CTCTTATGAC
 901 CCTTCAGACA CGTAGAGACA AGTACGATAC AAGTCAGCTG ACGCTCTCGT CGTTAGTTGC
 961 GCTGTGCGCG GATTGTTTCA GCTTTTGTTC GGAAAGGACC GCGAGAGTTC TGGACAATCA
 1021 AAGCCGCGCA CCAAAAGGAT GGGAAAAATCA GTTTGCGGCG GTTGCCAAAC AACGTCGACA
 1081 AGACCACTGT GAACTTTGCG AACGTTACTG CCATCGGGTT TCATCTGTTT GCCCTGCCTC
 1141 GAAGGATATC GAACGTTTTG AGGCTTTGGG ATATAAGGAG GTAGCACGAA GGACTTCCTT
 1201 AGATGGATTA CCCACGGTTC AGCAGATGTC TAGGTGCGAG AACGCTGTTT CGGAGTTCTC
 1261 AAGAGTTCTT TTTTCACTTG CGCAAAAGCC TAGTAAGGTG GCATTGTAAG ATGTTGACCC
 1321 GAAAAGTGGC TTGAAAAAAA GTGAGGAAGG TGGTGATGGT GAGCCAAACG AGCTTCATAA
 1381 AAAGGTACTT AGTGGTGGAT CAATAGATCA CGCCATCCTC CAGCGTTGGA CCACTGCTAG
 1441 GAAGGGACGA GCCGAAATGG AGGAGATGG AGTGATTGCA AGTGCGGAAG AGACCGCATC
 1501 GCTCAGGAG GAACTTGAC GTACAGGTC CGAAGATAC ATTAAGACAA GGGGACGAC
 1561 TCCGAAGTGG TGAGCAAGCA CGAAGAAGCA CAAACAAGC AAAACAACAA AAGAACAACA
 1621 AGTTTATTTT TTATTTTCT CTCTCTCTT TTTTGTGTC TCTCTCTAT ATCTTGGTAG
 1681 CAACTCTTTT TTCCCCCTC TCTTCTATA TATTAATTTT TCATTTGTTA GTTTTCTTT
 1741 ACCATGCGAG GTAGTTTAC GCGAGATGTG CCGATAGTTT TTTTATTATC GATCGATATG
 1801 GTAGCCGGAT GTGTGGCGGG AAAGACTTAT TTACTTCACG TTAACCGGG AGACAAGATT
 1861 CACGCGTCTG CGCTCCGCT CATTCAGAC GATGGAGGAA AGCCTCAGT GAACCTGTCT
 1921 CGCATCTTTA CGTCAACGCT GCAACGACAG AACGATTGGA AGTACAAGAA GAAGCCACTC
 1981 AAGGTTCCGA TTGAAGATT CGCCTCCCAT TTTGTGACAA GTTTGTCTTA CGACTTGAGC
 2041 AAGAACATTG TTGAAGAAGC TTTCCGGAG AGAAGGTGA TTATCCGTGA ACGGTGACGG
 2101 TTAACAGAGG TTCATTTAAG ACTGAGGGGT GTAAAGAAA GTCTTTATAT CCGCAACAA
 2161 GATGTTCAAA CTGC

TcP2β

AGP2β-1

AGP2β-2

1 TATTCTGCTT TATAAGTTGT TTTGACAGAT GAAGTACCTT GCGCTTACG CTCTTGTGTC
 121 CCTCTCAGGG GGGACCCCTT CAAAGTCTGC AGTTGAGGCC GTATTGAAGG CTGCCGCGT
 181 TCCTGTCGAT TCTTCTCGCG TGGATGCACT TTTTGCCGAG TTTGCAGGCA AGGACTTTGA
 241 TACGGTATGC GCTGAAGGCA AATCGAAGCT TGTAGGAGGG GTTGCACGTC CCAATGCTGC
 301 CACCGCTTCT CGACCCACTG CGCTGCCGC CGCTCTGGT GGTGCCGCG CGCTGCTGC
 361 AGCTGCGGAA GAGGAGGAGG ACGATGACAT GGGATTTGGT CTTTTGTACT AAAGACAAAT
 421 GCTTATTTTA TTTTATTTT GATTTTGTTC CTCTGTTGTC CATGAAGAGG CGCTGGAGTC
 481 CTGTGAATAA AATGCGAGAG AGTAAATAAA AGTAGGAACA GTAACGTCTG CACTCAAACG
 541 CATGGCAAAA AAGGGTACGC CATTTGGGTTG GCCTTTTTTT TCTTTTTTTT TCCCTTTTAA
 601 TTTTGTCTCT TATGACCCTT CAGACACGTA GAGACAAGTA CGATACAAGC CAGCTGACGC
 661 TCTGCTGCTT AGTTGCACTG TCGCGGGATT GTTCGAGCTT TTGTTTGGAA AGGACCGCGA
 721 GAGTTCTGGA CAATCAAAGC CGCGACCAA AAGGATGGGA AAATCAGTTT GCGCGGTTG
 781 CCAACAACG TCGACAAGAG GACTGCGAAC TTTGCGAACG TTAAGTCCAT CGAGTTTCAT
 841 CTGTTTGCCC TGACTCGAAG GATATCGAAC GTTTTGAGGC TTTGGGATAT AAGGAGGTAG
 901 CAAGAAGGAC TTCCTTAGAT GGATTACCCA CGGTTTCAGTA GATGCTAGG TGCGAGAAGC
 961 CTGTTTCCGA GTTCTCAAGA GTCTTTTTTT CACTCGCGCA CAAGCCTAGT AAGGTGGCAT
 1021 TTGAGGATGT TGACCCGAAA ACTGGCTTGA AAAAAAGTGA GGAAGGTGGT GATGTTGAGC
 1081 CAACCGAGCT TCATAAAAA GTACTTAGTG GTGGATCAAT AGATCACGCC ATCTGACGAG
 1141 GTTGAGACGAC TGCTAGGAAG GGACGAGCCG AAATGGAGGA GATGGAGTG ATTCGAAGTG
 1201 CGGAAGAGAC CGCATCGCTC AGGCAGGAAC TTGGACGTCA CAGGTCCGAA GAGTACATTA
 1261 AGACAAGGGG GACGACTCCG AAGTGGTGA GCAAGCAGAA GAAGCAGAAA CAAAGCAAAA
 1321 CACAACAAGA ACAACAAGTT TTATTTTAT TTTTCTCTCT CTCTCTCTT TTTTGTGCT
 1381 CTCTCTCTAT ATCTTGGTAG CAACCTTTT TTCCCCCTC TCTTTCTATA TATAATTTT
 1441 CATTGTGTAG TTTTCTTTA CCATGCAGAG TAAGGTTACG CGAGATGTGC GCATAGTTT
 1501 TTTTATATCG ATCGATATGG TAGCCGGATG TGTGGCGGGA AAGACTTATT TACTTCACGT
 1561 TAAACGGGGA GACAAGATTC ACGGCTCTGC TCCTCCGCTC ATTGCAGACG ATGGAGGAAA
 1621 GCTTCACGTG AGCTTGCTC GCATCTTTAC GTCAACGCTG CAACGACAGA ACGATTGGAA
 1681 GTACAAGAA AAGCTACGCA AGGTTGCGAT TGAAGAGTTG CGCTCCCAT TCGTGACAAG
 1741 TTTGCTCTAC GACTTGAGCA AGAACATTGT TGAAGAACT TTTCCGGAGA GAAGGTGAT
 1801 TATGCGTGAA CGGGACGGGT TAACAGAGGT TCATTTAAGA CTGAGGGGTG TAAAGAAAAG
 1861 TCTTTACATC CGGCAACAA AGTTTCAAAC TGCTTCCAAT GCATCTCTA GCTCTACAT
 1921 GCGCGACGCG AGCACCGCT CTTGTTTGG ACCCCACGAT CCCTCTGAGG AACATATTAA
 1981 TCATTCAACC ACTTCATTG TGGTTGACGA TACTTTAAGC TCATTTGATA GTTTGACAGG
 2041 GACTGCAAGC GAGCAAAGGA GCGAGGACAT CCATTCTGAG ACGAAAGAAA CTCAGGTGGA

TcP2β

AGP2β-1

AGP2β-2

Figura 31-
 Secuencia del frag-
 mento genómico G9,
 que contiene al locus
 TcP2β-H1.5. Sub-
 rayado se señalan las
 regiones de polipirimi-
 dina involucradas en
 trans-splicing. En
 negrita y subrayado se
 marcan los codones de
 iniciación y terminación
 de los genes.

Figura 32-
 Secuencia del frag-
 mento genómico G8,
 conteniendo al locus
 TcP2β-H1.3. Sub-
 rayado se señalan las
 regiones de polipirimi-
 dina involucradas en
 trans-splicing. En
 negrita y subrayado se
 marcan los codones de
 iniciación y terminación
 de los genes. Obsérvese
 que en AGP2β-1 el
 aceptor AG está mu-
 tado y probablemente
 en este locus no se ex-
 prese.


```

2101 CAAAGAACTC AATAAAGAA ATTTTGACAA GTGCGCTTCA GCGCCGGATA CATCTTTACA
2161 CCGGTTAGTT GCACCCATGC ATCGAAGGAC GGGGGGCGTG TCAAATACAA TCAGCAAGGA
2221 AAACTTTAAG AATCTTGCTT TTACACGGGG CTCAGATGCG GCTATTTTGA AGGGTAAGGA
2281 TTTATCCACT CTTAAGCCGT CTTGGAAGAG AAATATTGCT TTTGAGGACC CCGTTCATGG
2341 GGTCATTAGG CAGATTTGTG ATGCTGTAAG GTGGTGTGGA GACATTGGTA GGGAGTCTCA
2401 AACAGCGAGA GATGCTTCGA ATGGAGCTCA ATTAGCTCTT GATTTTTATT TAAATCAGGA
2461 TGGAGAAAAC CTGGGTCGCT TTGTTGTTAG TTTTCTGGCT CATGTTGTTG TTGAAGTGTG
2521 TCGAGAGACG CGACACGTTG GCAACTGGCT TAGCCTTTTA ACACACGTTA TTTACGGATC
2581 C

```

dinucleótidos aceptores AG (ver figura 29). Este tipo de organización indica que muy probablemente estos genes se transcriban.

La secuencia aminoacídica derivada de AGP2 β -1 codifica un polipéptido de 225 aminoácidos ($pI = 7.82$), mientras que la secuencia derivada de AGP2 β -2 esta incompleta y probablemente se extienda mas allá de los límites del fragmento genómico clonado. G8 contiene la versión más larga de la proteína, 373 aminoácidos. Las secuencias proteicas (figura 34) fueron comparadas en la base de datos SWISS-PROT (última comparación, Agosto de 1995). No tienen homología con ninguna proteína conocida hasta la fecha. Una búsqueda en PROSITE (base de datos de motivos proteicos) encontró un motivo consenso de unión de ATP/GTP en AGP2 β -1, AGC-VAGKT.

ENTRY AGP2 β -1 #Type Protein
SUMMARY #Molecular-weight 25307 #Length 225 #Checksum 406

```

1 M T L Q T R R D K Y D T S Q L T L S S L V A L C A D C S S F
31 C L E R T A R V L D N Q S R A P K G W E N Q F A A V A K Q R
61 R Q E D C E L C E R Y C H R V S S V C P A S K D I E L F E A
91 L G Y K E V A K E R T S L D G L P T V Q Q M S R C E N A V S E
121 F S R V L F S L A H K P S K V A F E D V D P K S G L K K S E
151 E G G D G E P T E L H K K V L S G G S I D H A I L Q R W S T
181 A R K G R A E M E E M G V L R S A E E T A S L R Q E L G R Q
211 R S E E Y I K T R G T T P K W STOP

```

ENTRY AGP2 β -2 #Type Protein
SUMMARY #Molecular-weight 41939 #Length 373 #Checksum 3709

```

1 M Q S K V T R D V R I V F F L S I D M V A G C V A G K T Y L
31 L H V K R G D K I H A S A P P L I A D D G G K L H V S L S R
61 I F T S T L Q R Q N D W K Y K K K L R K V R I E E L R S H F
91 V T S L S Y D L S K N I V E E T F P E R R V I M R E R D G L
121 T E V H L R L R G V K E S L Y I R Q Q D V Q T A S N A S S S
151 S Y M R D A S T A S L F G P H D P S E E H I N H S T T S F V
181 V D D Y L S V I D S L T G T A S E Q R S E D I H S E T K E T
211 Q V D K E L N K R N F D K C A S A P D T S L H P L V A P M H
241 R R T G G V S N T I S K E N F K N L A F T R G S D A A I L K
271 G K D L S T L K P S W K R N I A F E D P V H G V I R Q I C D
301 A V R W C G D I G R E S Q T A R D A S N G A Q L A L D F Y L
331 N Q D G E N L G R F V V S F L A H V V V E V C R E T R H V G
361 N W L S L L T H V I Y G S ...

```

Figura 34-

Secuencia de aminoácidos de los genes asociados a TcP2 β , AGP2 β -1 y AGP2 β -2. Los puntos indican que la secuencia continua más allá de los límites de clonado. En negrita se marcan las cisteínas repetidas (ver pag. 85)

```

H6.4  GGAATTCCTCTCTAGGAGAAATTTTGGACACACGATTAGACAATAGTTTGGATGAGTGCAAGACCATAGGTGACGTAAATGCCCTATTCTCTGAAGCGGATT
H1.8  .....A.....T.....
H1.5  .....A.....T.....

H6.4  GCCTACCTTGCCCTAACACGAGAGCGGTTAACGAGTTTCTTCAAAATATGCAGCGGATTCACTAAGAACATTTTCACGCACGAAAGCGAAATTATTATGATTGTTATTATATACITTTTCTTTGTT
H1.8  .....A.....
H1.5  .....C.....C.....

H6.4  GTTTTATCCACTTATTATTGTTGTTGTTAAATGTTTACCTTTTCTTTTCCCACTCTTTTATGATGCTCTTTCTTTT.....TTGCTTTATAACTTGTCTTGTGTC
H1.8  .....T.....A.G.....CTT-----SIRE-----TTATTC.....
H1.5  .....T.....A.G.....SIRE-----TTATTC.....T..A.
H1.3  .....SIRE-----TTATTC.....

H6.4  AGATGAAGTACCTTGCCCGTTACGCTCTTGTGGCCTCTCAGGGGGGACCCCTCAAAGTCTGCAGTTGAGGCCGTATTGAAGGTCGCCGGTGTTCCTGTCGATCCTTCGCGTGGATGCACCTTTTACC
H1.8  .....T.....C.....G..
H1.5  .....A.....C.C.....G..
H1.3  .....C.....T.....G..

H6.4  GAGTTTGGGGGCAAGGACTTTGATACGGTATGCACCTGAAGGCAATCGAAACTTGTAGAGGGGTACACGTCCCAATGCTGCCACCGCTTCTGCACCCACTGCCGCTGCCGCCCTCTAGTGGTGCCTGC
H1.8  .....C.....G.....
H1.5  .....T.A.....G.....*
H1.3  .....A.....G.....*

H6.4  CGCGCTGCTGCTCAACTGCGAAGAGGAGGAGGACGATGACATGGGATTTGGTCTTTTGTGACTAA AGACAAATGCTTATTATTATTATTATTGTTTCTCTGTTGTCCATGAGAGGCGCTGGAG
H1.8  .....G.....T.....C.....
H1.5  .....G.....T.....A.....
H1.3  .....G.....|.....

H6.4  TCCTGTGAATAAAATGCGAGAGAGTAAATAAAGTGGGAACAGTAAGTCTGCTCAACTCAAAAGGATGGCAAAAGGATGGGTTGGCCTTTT-----SIRE-----TcP2β-----
H1.8  .....A..T.....TT-----SIRE-----TcP2β-----
H1.5  .....A.....
H1.3  .....A.....

H6.4  -----TTTTCCCTTTTAAGTTTGTCTT ATGACCCCTCAG
H1.8  -----SIRE-----TTATCTCTT.....C.....
H1.5  -----CCC.....C.....
H1.3  -----TCTTT.....T.....C.....

```

Figura 33-

Comparación de los loci TcP2β.

Los puntos representan bases idénticas. Los cambios de base se indican con respecto a TcP2β-H6.4. Las líneas punteadas fueron introducidas para mantener el alineamiento y muestran la inserción de una secuencia (SIRE) en H1.8 y H1.3 que no está en los otros loci.

Subrayado y en **negrita** se muestran los codones de iniciación y terminación de los genes TcP2β. Igualmente pero en *itálica*, se muestran las señales de trans-splicing río arriba de TcP2β. Las estrellas (*) indican los distintos sitios de poliadenilación encontrados en los transcritos de los loci TcP2β.

Sin embargo, un análisis detallado de la proteína AGP2 β -1 identificó un motivo de cisteínas espaciadas regularmente, típico de los factores de transcripción Zinc-fingers:

C X2 C X3 C X33 C X2 C X3 C X6 C

La letra X puede ser cualquier aminoácido y el número indica la cantidad de residuos que separa una Cys de la otra. Como en el caso de las proteínas Zinc-finger estos motivos se encuentran rodeado por aminoácidos básicos. Un motivo similar se encuentra en el gen ESAG-8 de *Trypanosoma brucei* (72). Si bien en tripanosomas no se han clonado aun factores de transcripción la presencia de este motivo es llamativa, pero no existen otros elementos de juicio para afirmar que se trate de un factor de transcripción.

Otro hecho notable es la compactación de los 3 genes en estos loci cromosómicos (ver figuras 28 y 29), las regiones intergénicas que los

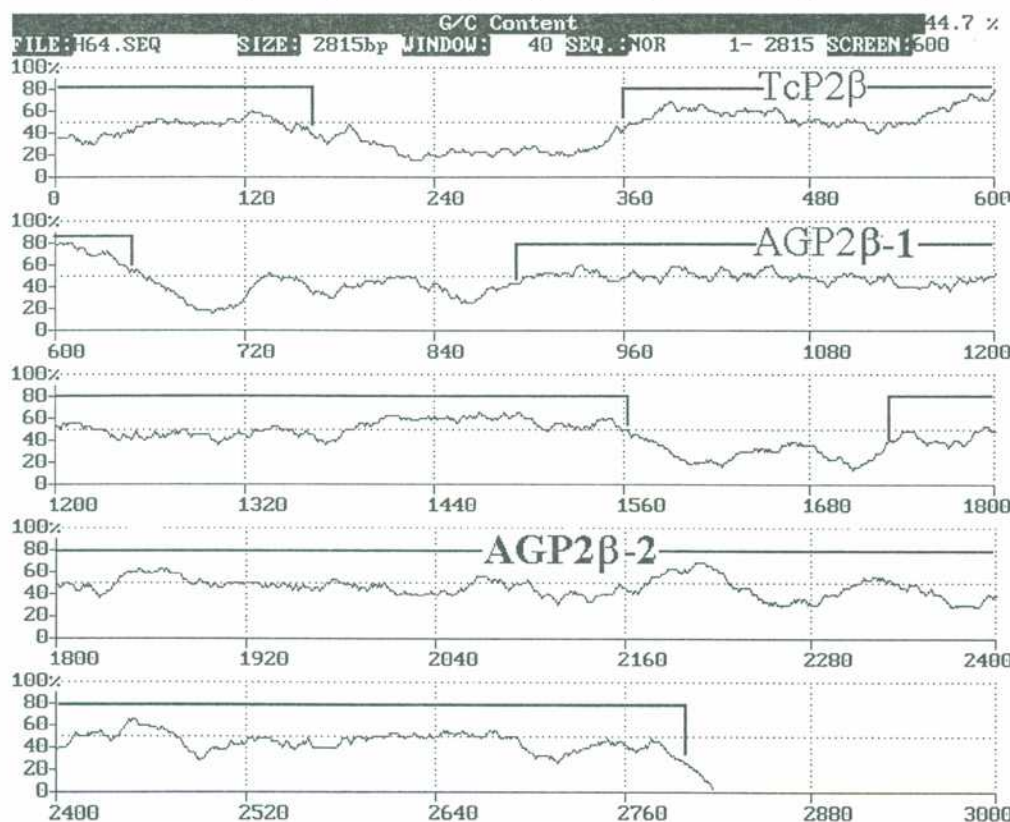


Figura 35-

Diagrama del contenido en GC del locus H6.4 generado con el programa DNASIS. El porcentaje indica el contenido de GC. La media es de 44.7% para todo el locus. Obsérvese la caída del contenido en GC en las zonas intergénicas.

separan son de naturaleza repetitiva, ricas en T, y de muy bajo contenido en G+C (ver perfil G+C en figura 35). Estos resultados concuerdan con los datos obtenidos para otros loci de tripanosomas (33, 34, 35, 73). Este tipo de organización es consistente con la transcripción policistónica de estos parásitos (48). El bajo contenido en G+C de la zonas intergénicas podrá ser de utilidad marcadora de marcos abiertos de lectura, necesidad urgente de los proyecto genoma de parásitos. (figura 35).

4.2- Organización de los loci TcP2 β .

El análisis por Southern blot muestra que efectivamente los genes TcP2 β se encuentran distribuidos en cuatro loci genómicos. La hibridización con la sonda 3' TcP2 β del ADN genómico digerido por HindIII da 4 bandas : 1.3 kb; 1.5 kb; 1.8 kb y 6.4 kb. El locus H6.4 carece del sitio HindIII interno en TcP2 β (ver figuras 28 y 29). En cambio, la hibridación con el ADNc completo revela 7 bandas. 3 de ellas corresponden a las regiones 5' de los loci H1.8, H1.5 y H1.3 (figura 24, calles 5 y 6, capítulo 3).

Es interesante observar que la intensidades de las cuatro bandas Hind III (1.3; 1.5; 1.8; 6.4 kb) (figura 24, calle 6), no son iguales. Esto se verificó no solo para los cortes HindIII, sino también para el resto de las enzimas . Por densitometría sobre las autorradiografías, se puede observar que los fragmento originados por corte con EcoRI también presentan diferencias en la intensidad de hibridación:

| PICO | BANDA | AREA RELATIVA
(%) | NUMERO DE
COPIAS |
|------|--------------|----------------------|------------------------|
| 1 | EcoRI 1 Kb | 11.1 | 2 |
| 8 | EcoRI 5.6 Kb | 14.5 | 2 |
| 9 | EcoRI 7 Kb | 28.2 | 4 |
| 10 | EcoRI 10 Kb | 14.5 | 2 |
| | | | TOTAL: 10/
DIPLOIDE |

Esto significa que habría 5 copias del gen TcP2 β por genoma haploide de

T. cruzi, resultado consistente con el mapeo mostrado en la figura 28.

Los cuatro loci clonados son altamente homólogos a nivel de secuencia. Esto indicaba que las regiones clonadas no contienen puntos de ruptura (breakpoints). Por puntos de ruptura entendemos aquellas regiones 5' y 3' en donde los loci provenientes de cromosomas no homólogos empiezan a diferir en secuencia. Estos puntos indican también las regiones que delimitan la zona duplicada y translocada en el curso de la evolución para generar una nueva copia génica.

El análisis de los puntos de ruptura se realizó por southern blot con sondas diferenciales 5' y 3' obtenidas de los loci recombinantes TcP2 β . Estas son, PH500, derivada de la zona intergénica entre AGP2 β -1 y AGP2 β -2 y BH950, derivada de la zona codificante de AGP2 β -2 (figuras 29 y 35). Como control se usó la sonda TcP2 β . Este análisis de southern blot demuestra lo siguiente:

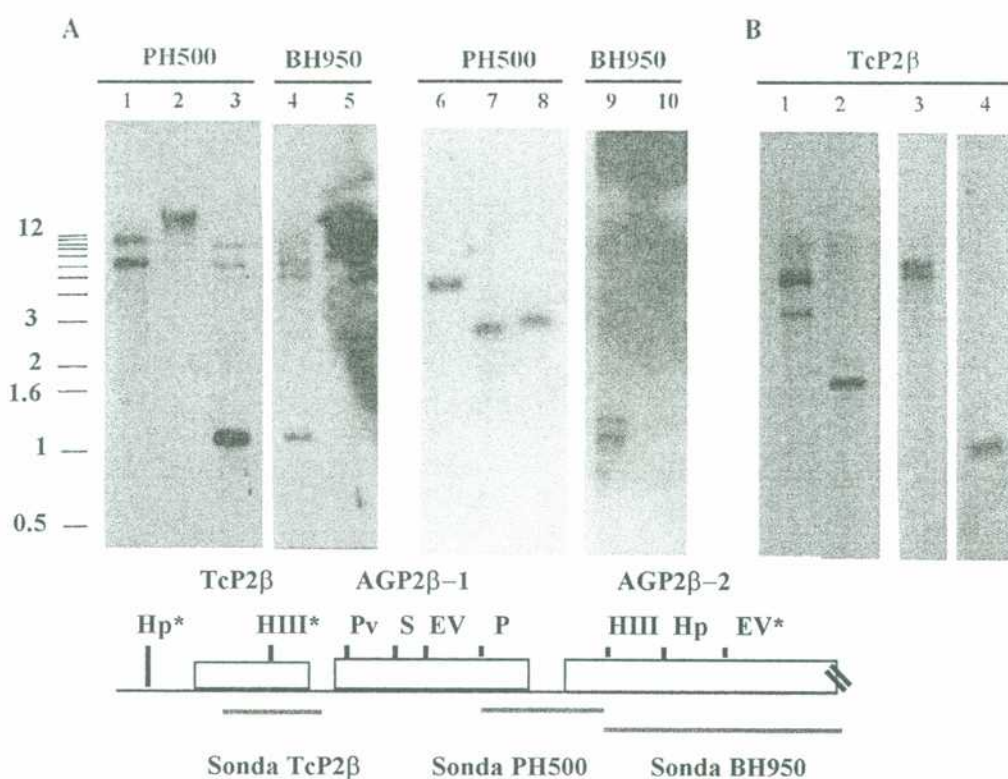


Figura 36-

Southern blot de ADN genómico. **A.** 1, Sall. 2, Sin cortar. 3 y 4, EcoRV. 5, HindIII. 6, Sall+EcoRI. 7, PvuII. 8, PstI. 9, BamHI+HindIII. 10, HindIII. **B.** 1, Sall. 2, HpaI. 3, EcoRV. 4, Sall+EcoRI.

En el diagrama inferior se muestra la ubicación de las sondas. Hp, HpaI. HIII, HindIII. Pv, PvuII. S, Sall. EV, EcoRV. P, PstI. El asterisco (*) indica que el sitio es polimórfico en algunos loci.

Los marcadores de PM son 1Kb Ladder (GIBCO, BRL.)

La hibridización con la sonda PH500 muestra una sola banda con las enzimas PvuII, PstI y Sall + EcoRI (figura 36 A., calles 6, 7 y 8), indicando que los dos primeros sitios y EcoRI se conservan en todos los loci. En cambio los sitios Sall y EcoRV son polimorficos (figura 36. A, calles 1 y 3). De acuerdo a los tamaños de las bandas hibridizantes con PH500, el resultado indica que hasta aproximadamente 6 Kb rio abajo de la sonda los loci mantienen homología.

La hibridización con las sonda BH950 detecta solamente 2 bandas HindIII (figura 36 A. calle 5). El análisis del corte EcoRV (calle 4) es consistente con los datos de PH500 indicando que ese sitio es polimórfico dentro de la zona clonada (sitio interno de AGP2 β -2)

El resultado de ambas sondas indica que los loci conservan homología al menos hasta 6 kb rio abajo del gen TcP2 β , después de esa distancia solo se detectan dos bandas (por ejemplo en los cortes Sall con PH500 y HindIII con BH950)

Asi, existen en realidad dos tipos diferentes de loci, que podemos denominar aquí Tipo I y Tipo II. Los otros dos loci derivan de estos por la inserción de una secuencia de 420 pb (ver detalles de esta secuencia mas adelante). La observación de 4 bandas en southern blot con EcoRI, KpnI, BamHI y Sall hibridizados con sonda TcP2 β (figura 24) son explicables por esta inserción.

Estos dos tipos de loci mantienen una zona mínima conservada entre ambos de aproximadamente 6 kb. Los fragmentos genómicos clonados quedan incluidos en esta región. Se puede deducir del mapeo que el punto de ruptura 3' se encontraría de 4.5 a 6.5 kb hacia abajo del gen TcP2 β y un punto de ruptura 5' se encontraría mas cerca, entre 1 y 2 kb hacia arriba del gen TcP2 β (dado que el sitio EcoRI 5' que se ve en los mapas de la figura 28 es el ultimo sitio en común que se observa en estos mapeos por southern, figura 36 B., calles 1-4).

Cuando una fragmento genómico se duplica y transloca generando una

nueva copia generalmente arrastra consigo el promotor de la unidad transcripcional, dado que sino dependería en gran medida de la presencia azarosa de un promotor correctamente posicionado en el nuevo lugar de translocación. De aquí la importancia que el punto de ruptura 5' se encuentre tan cerca del gen TcP2 β , ya que es muy probable que el promotor de toda la unidad transcripcional se encuentre dentro de ese rango de 1 a 2 kb.

El kariotipo electroforético de T.cruzi CL Brenner muestra que 6 bandas cromosómicas hibridizan con TcP2 β : IX, X, XII, XIV, XV, XVI y XIX según la nomenclatura que proponemos (74) (dato no mostrado). Es probable que algunos de estos cromosomas sean pares homólogos, dado que solo se identifican 4 loci usando la técnica de southern blot. Que los pares homólogos no tengan el mismo tamaño por PFGE, puede deberse al polimorfismo cromosómico originado del alargamiento o acortamiento en las regiones teloméricas durante el ciclo de división (74).

4.3- Polimorfismo de los loci TcP2 β . Análisis por SSCA.

La organización de estos loci se mantiene en las cepas Y, CL y Maracay, pero no en la cepa G, cepa aislada de comadrejas. En esta cepa los resultados son consistentes con la existencia de un único locus para TcP2 β (69).

Un análisis mas detallado se puede hacer por la técnica SSCA (Single Strand Conformation Analysis, análisis de conformación de cadena simple).

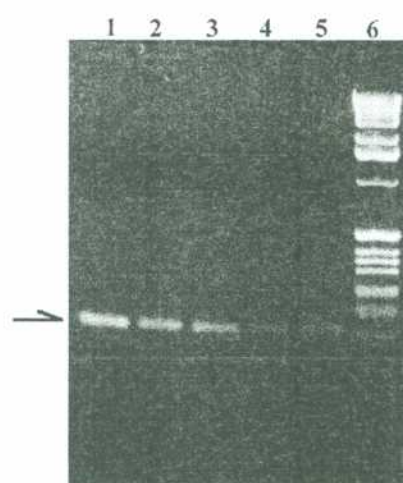
Este análisis ha sido usado para detectar mutaciones puntuales en enfermedades genéticas (75). El procedimiento requiere una amplificación por PCR de una secuencia de aproximadamente 100 a 500 pb y se basa en el principio que la cadena simple adquiere una estructura secundaria dependiente de la secuencia nucleotídica cuando se corre en geles no desnaturizantes (75, 76). Si existen diferencias entre las secuencias que se comparan, se debe observar una migración diferencial de las cadenas.

La principal ventaja del análisis SSCA es su simplicidad y relativa sensibilidad. El conocimiento de la secuencia de los genes TcP2 β permite probar la técnica para esta familia de genes y determinar si puede ser usada para analizar la compleja organización multigénica de estos parásitos protozoarios. En estos organismos, los genes, que codifican para antígenos contienen motivos repetidos. En algunos casos como HSP70, copias idénticas del gen original se encuentran organizadas en tandem definiendo un locus repetitivo en un cromosoma (35). En otros casos como gp90, varias formas del mismo gen se encuentran dispersas por el genoma mapeando en diferentes cromosomas.

La familia génica TcP2 β puede ser, entonces, tomada como referencia para probar la eficiencia del SSCA para detectar heterogeneidad de secuencias en este parásito.

Un análisis detallado permitió identificar una región 5' de 56 pb donde 3 de los 4 loci clonados, H6.4, H1.8, H1.5, difieren en una sola base entre ellos (figura 37). El gen TcP2 β H6.4 se tomó como referencia, y las sustituciones fueron referidas a esta forma. Así, H1.8 presenta una única sustitución, una transición C a T, con respecto a H6.4, mientras que con H1.5 hay una transición G a A. Los tres loci clonados y el ADN genómico de las cepas Tulahuen y G se usaron como templados para amplificar la secuencia de 56 pb usando los cebadores indicados en la figura 37, para dar un amplicón de 97 pb (figura 38).

Figura 38-
Amplificación radiactiva de la secuencia 5' de los genes TcP2 β . (97 pb). Se usó como templado ADN de las siguientes fuentes. 1, H6.4. 2, H1.8. 3, H1.5. 4, cepa G. 5, cepa Tulahuen. En la calle 6, se corrió 1 kb ladder como marcador de PM.



Se probaron cuatro condiciones electroforéticas distintas para determinar cual era la mejor para identificar las variantes. Los experimentos se llevaron a cabo en geles 6% de poliacrilamida con o sin el agregado de glicerol corridos temperatura ambiente o a 4°C (ver materiales y métodos, pag...)

Como se muestra en la figura 39 A., los geles corridos sin glicerol a 4°C dieron la

| | S-1 | 50 | S-2 | 100 | 130 |
|------|--|-----|-----|-----|---|
| H6.4 | TTGCTTTATATAAGTTGTCTTGTGCAGATGAAGTACCTTGCCGCTTACGCTCTTGTGGCTCTCAGGGGGGACCCCTCAAAGTCTGCA | | | | TTGAGGCCGTTATTGAAGGCTTCCCGGTGTTCTCGATCCCTTC |
| H1.8 | C.....T..... | | | |C..... |
| H1.5 |A..... | | | |C..... |
| H1.3 | C.....T.....A..... | | | |C.....T..... |
| | | 180 | | 230 | 260 |
| H6.4 | TCGCGTGGATGCACCTTTTACCGAGTTTTCGCGGCAAGGACTTTTGATACGGTATGCACCTGAAGCAATCGAACTTGTAGAGGGGTACACGTCCCAATGCTGCCACGCTTCTGCACCCACTGCCGCTG | | | | |
| H1.8 |G.....A.....C..... | | | |C..... |
| H1.5 |G.....T.A..... | | | |G..... |
| H1.3 |G.....A.....G..... | | | |G..... |
| | | 310 | S-3 | 360 | 390 |
| H6.4 | CCCGCCGCTCTAGTGGTGC CGCGCGCTGCTGCAACTGCGGAAGAGGAGGAGGACGATGGGATTTGGTCTTTTGGACTATAGACAAATGCTTATTTTATTTTATTTTCTCTGT | | | | |
| H1.8 |G.....G..... | | | | |
| H1.5 |G.....G..... | | | | |
| H1.3 |G.....G..... | | | | |

Figura 37-

Detalle de la secuencia y los cebadores utilizados para amplificar los fragmentos analizados por SSCA. Los puntos indican identidad de bases. Los cambios se muestran respecto de H6.4. Recuadrado se señala la región utilizada como cebador. S-1 es sentido y S-2 y S-3 son antisentido

mejor resolución. Esta condición diferencio claramente la movilidad de los distintos amplicones. Ambas cadenas de H6.4 migraron mas rápido que las de H1.8; mientras que las de H1.5 presentaron un patrón distinto, una de las cadenas migró mas rápido que las de H6.4 y H1.8, y la otra fue la mas lenta (figura 39 A.).

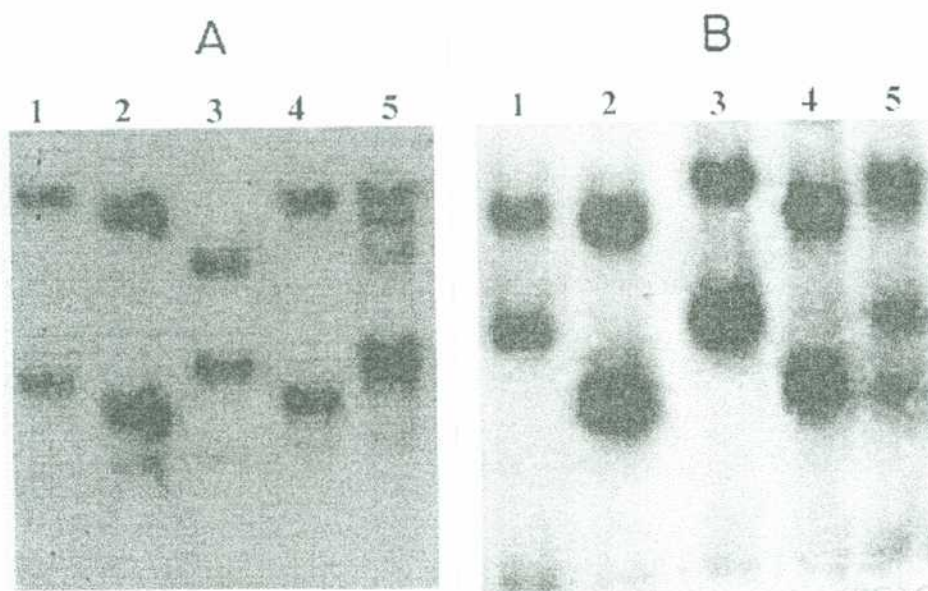


Figura 39-

Análisis SSCA del fragmento de 97 pb. Corrida sin la presencia de glicerol. A. Gel corrido a 4°C durante 7 Hs. B. Gel corrido a temperatura ambiente durante 7 Hs. 1, H1.8. 2, H6.4. 3, H1.5. 4, Cepa G. 5, Cepa Tulahuen.

Cuando se analizan los productos de amplificación del ADN genómico la cepa Tulahuen, se encuentra un patrón consistente con la combinación de las tres variantes clonadas. Sin embargo, cuando se analizo la cepa G, presento un patrón de migración consistente con la existencia de un solo gen que migra, en este caso, igual que la variante H6.4. La secuencia directa del producto de PCR demuestra que efectivamente la region 5' del gen de la cepa G es igual que H6.4 como predijo el SSCA.

Cuando el análisis se realiza en una corrida a temperatura ambiente los resultados anteriores se confirman. Nuevamente H1.5 presenta una migración diferencial (figura 39 B.).

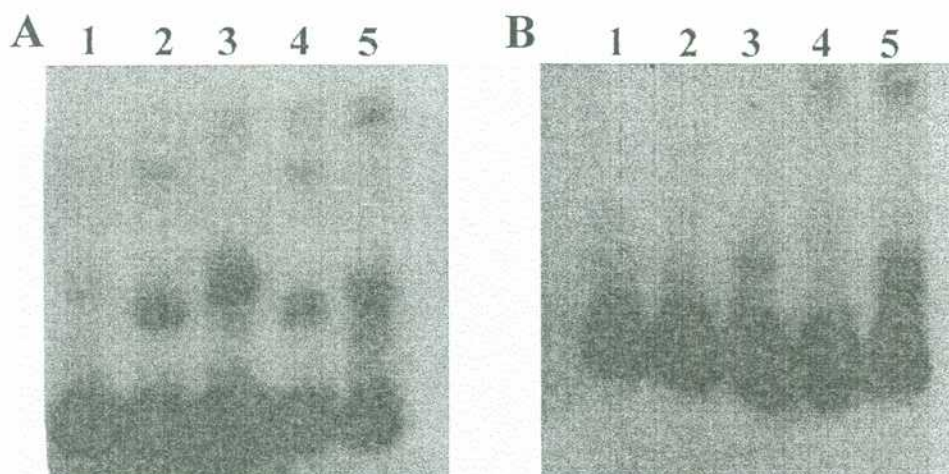
La figura 40 presenta los resultados de corridas en presencia de glicerol. En ningún caso la presencia de glicerol logró dar una separación adecuada

de las variantes. Solo H1.5 siguió diferenciándose.

Figura 39-

Análisis SSCA del fragmento de 97 pb. Gel corrido en presencia de 10% de glicerol. A, gel corrido a 4°C. B, gel corrido a temperatura ambiente.

1, H1.8. 2, H6.4. 3, H1.5. 4, cepa G. 5, cepa Tulahuen



Análisis de la secuencia codificante completa de TcP2β por SSCA.

La secuencia codificante completa de TcP2β es de 321 pb. Cuando las tres variantes estudiadas se comparan es evidente que difieren en varias mutaciones puntuales: H1.8 presenta 8 cambios nucleotídicos, transiciones, con respecto a H6.4. El locus H1.5 presenta 5 transiciones en la mismas posiciones que H1.8 pero, además, dos transiciones y una transversión características de esta variante. La cuarta forma TcP2β H1.3, presenta las mismas 5 transiciones más 4 transiciones específicas (figura 37)

Usando los cebadores S1 y S2 (figura 37) se amplificaron las 321 pb y se analizaron con la amplificación del ADN genómico de las cepas Tulahuen, Maracay y G.

El resultado de dos corridas electroforéticas sin glicerol se muestra en la figura 41. Comparando las distintas amplificaciones se puede observar que las cepas Tulahuen y Maracay presentan amplicones consistentes con la existencia de 4 variantes confirmando los resultados anteriores. En estas dos cepas la migración de bandas es igual, indicando la probable identidad de las variantes. La cepa G presenta una sola variante, pero ya no migra como

H6.4. Esto implica que la región de homología entre estos genes se limita solo al extremo 5', y que probablemente sea una variante específica de la cepa G.

Se puede concluir que es importante probar todas las condiciones de corrida antes de elegir una, dado que algunas no funcionan, pero eso depende de la secuencia que se analiza (77). También es importante el tamaño del fragmento analizado. Cuando se usó el fragmento de 100 pb se obtuvieron mejores separaciones de las cadenas que cuando se uso el fragmento de 321 pb.

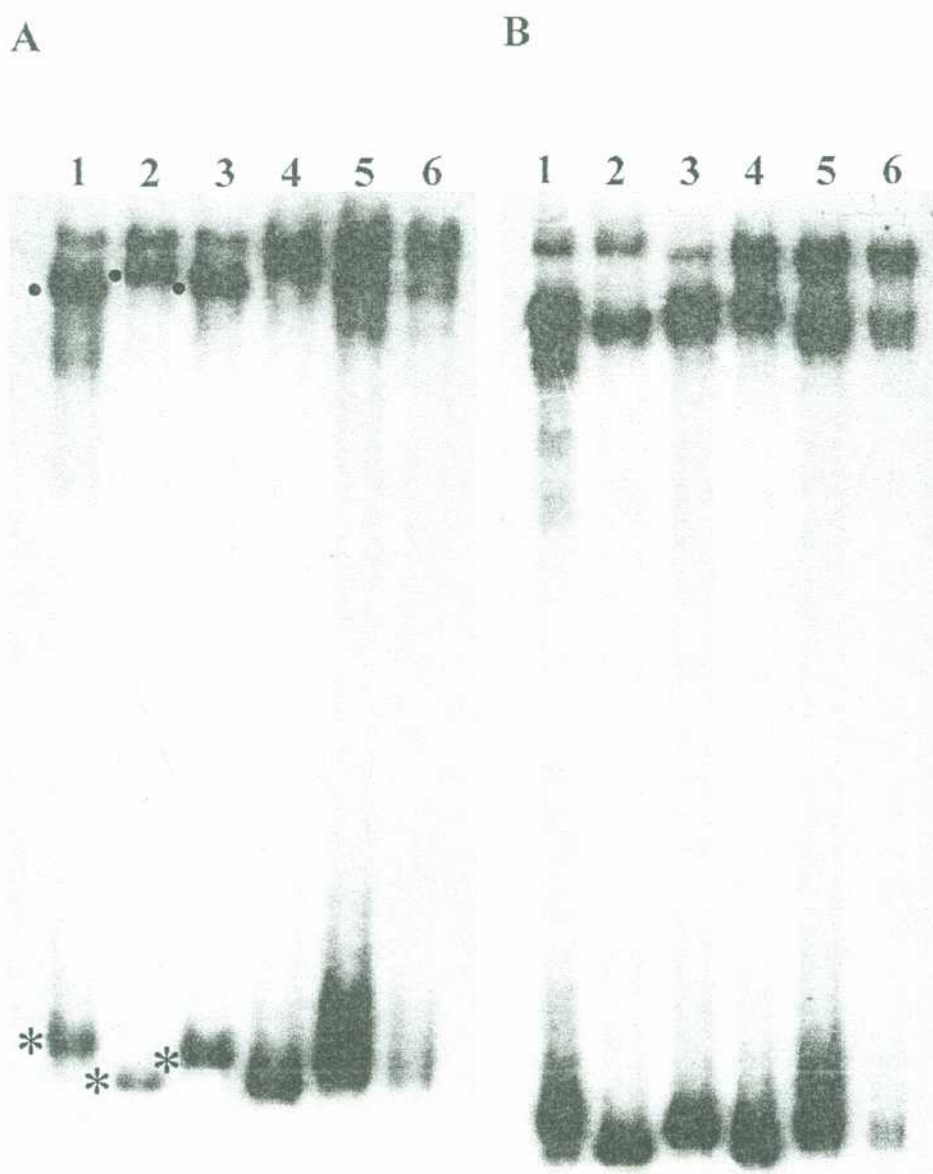


Figura 41-

Análisis SSCA del fragmento de 320 pb. Gel corrido sin glicerol. **A**, gel corrido a 4°C. **B**, gel corrido a temperatura ambiente.

1, H1.8. 2, H6.4. 3, H1.5. 4, cepa G. 5, cepa Tulahuen. 6, cepa Maracay. Los puntos y asteriscos muestran la separación de las variantes.

4.4- Descubrimiento de una secuencia repetida (SIRE) insertada en un locus TcP2 β : características y consecuencias sobre la expresión.

Uno de los hechos que se puede observar definitivamente luego de mapear los 4 loci, es el tandem de genes TcP2 β presentes en el locus H1.8. Estos dos genes mantienen un 100% de identidad a nivel de secuencia y están conectados físicamente por un fragmento de ADN de 428 pb que está ausente en los otros loci. Esta secuencia se encuentra también ubicada inmediatamente 5' del primer gen del tandem e inmediatamente 3' del segundo (ver la secuencia completa de este locus en figura 30).

La comparación con los otros tres loci claramente definió que esta secuencia de 428 pb se encontraba insertada en este locus interrumpiendo las regiones intergénicas en las tres posiciones indicadas. Una pequeña región 3' de esta secuencia también se encontró al comienzo del clon genómico correspondiente al locus H1.3, delante del gen TcP2 β .

Como esta secuencia se encuentra repetida y dispersa por el genoma (ver detalles mas adelante), se la denominó **SIRE** (Short Interspersed Repetitive Element).

La asociación de SIRE con TcP2 β resulta evidente de solo observar los datos de Southern blot. Los fragmentos de 1 kb EcoRI y BamHI se pueden predecir de la distribución de SIRE en el locus H1.8 (figuras 24 y 28) (nótese que SIRE tiene sitios únicos para EcoRI y BamHI). Además, se puede observar que la banda de 2.7 Kb BamHI de southern blots es compatible con la banda que se deduce del mapa del locus H1.3 que tiene una única inserción de SIRE (desde el sitio interno BamHI presente en SIRE hasta el sitio BamHI interno del gen AGP2 β -2) (figuras 24 y 28).

Caracterización de la secuencia SIRE insertada en TcP2 β .

La comparación de los cuatro loci determinó que la inserción de SIRE en H1.8 genera una duplicación de 5 bases con el siguiente consenso: TTATT. De hecho, el sitio de inserción del SIRE parece ser una región de **polidT** presente en la parte 5' de cada gen (una repetición de 12 a 14 T seguidas) (ver figura 30). Así esta secuencia TTATT parece provenir de la duplicación del polidT por inserción del SIRE. Es interesante notar que el retrotransposón TRS-1.4 de *Trypanosoma brucei* también genera la duplicación de 4 o 5 T en su lugar de inserción (78). El SIRE en si mismo se expande entre repeticiones directas de 4 bases (AGGA) presentes en su parte 5' y 3' justo antes de la zona de duplicación (figura 30).

Todas estas características son típicas de transposones y secuencias repetidas dispersas por el genoma. Sin embargo, SIRE no lleva la región de poli A típico de los retrotransposones eucarióticos.

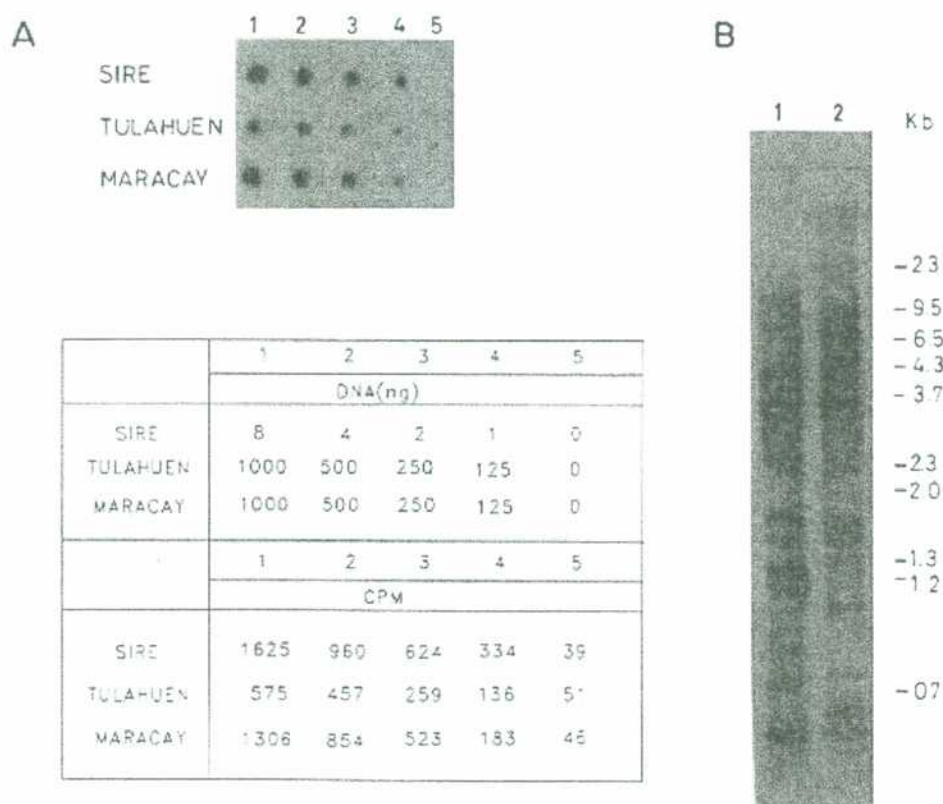
Observando la región 5' de SIRE se pudo encontrar una secuencia similar a la secuencia recombinógena **Chi** del fago λ , **GCTTGTG**, que ha sido identificada en otros Trypanosomas como una secuencia asociada a regiones que sufren recombinación (79) (figura 30 y 43). Esto es coincidente con la particular organización que muestra el locus H1.8 (figuras 28 y 30), en donde se puede observar la duplicación del gen TcP2 β de la siguiente forma: SIRE-TcP2 β -SIRE-TcP2 β -SIRE. La situación más simple de imaginar la generación de esta organización es proponiendo que la inserción de SIRE en 5' y 3' de un gen TcP2 β original, favoreció la recombinación entre los SIRE para producir la posterior duplicación del gen TcP2 β .

Otra característica común de estos elementos es su inserción en varios lugares del genoma, como consecuencia generando un número de copias elevado. Para verificar si este era el caso del SIRE, se realizó un Southern Blot genómico cortando ADN de *T. cruzi* cepa Tul II con EcoR I y BamH I e hibridizando con una sonda de la secuencia SIRE amplificada por PCR. Por

otra parte, se hibridizó un Dot Blot con diluciones seriadas al medio de ADN genómico y de ADN de la secuencia clonada (locus H1.8) con el objetivo de cuantificar el número de copias presentes (figura 42).

El Dot Blot determinó que existen aproximadamente 2400 copias de SIRE en la cepa Tul II y 4400 en la cepa Maracay asumiendo un tamaño de genoma de $0.8 \cdot 10^8$ pb. El Southern Blot muestra que SIRE se encuentran disperso y en sitios diferentes.

Figura 42-
A, Dot blot de ADN genómico para cuantificar las copias de SIRE. En el panel inferior se muestran los datos de concentración y cpm correspondientes a cada punto. La secuencia clonada de SIRE se usó como referencia.
B, Southern blot genómico hibridizado con SIRE. 1, BamHI. 2, EcoRI.



El ligamiento de la secuencia SIRE y el gen TcP2 β se encuentra conservada en las cepas Tul II, Y, CL y Maracay (como se deduce de las bandas EcoRI 1 Kb y Hind III 1.8 Kb, por ejemplo, de los Southern Blots) (69), pero no en la cepa G, que es la única cepa mostrada que no deriva de un hospedador humano. En efecto, en esta cepa el único locus TcP2 β no se encuentra asociado con un elemento SIRE. Evidencias obtenidas de South-

ern blot usando las sondas TcP2 β y PH500 sobre ADN genómico de la cepa G muestran que se mantiene la organización general del locus de Tulahuen (TcP2 β ---AGP2 β -1---AGP2 β 2) pero sin la presencia de SIRE (figura 36).

Una búsqueda en las bases de datos GeneBank y EMBL con la secuencia SIRE mostró que esta secuencia es de naturaleza repetida: varias secuencias genómicas de *T. cruzi* informadas presentaron distintos grados de homología con SIRE. Por ejemplo, una zona de 400 pb presente en la región 3' no codificante de la Histona H2A (X67287) es 70 % homologa a SIRE, una región de 258 pb presente en el Pseudogen ribosomal a 24S también presentó homología con SIRE (80). Finalmente, las últimas 122 pb de SIRE resultan 95% homólogas a las primeras 122 pb del elemento repetitivo E13 1Kb descrito por Requena et al. (81) (figura 43). Sin embargo, la similitud entre estos dos elementos se limita solo a esa región.

Figura 43-

Comparación de secuencias de *T. cruzi* encontradas en la base de datos EBI con SIRE.

Los puntos indican bases iguales. Las líneas se usaron para mantener la homología. Recuadrada se señala la secuencia Chi y en negrita y subrayado las señales de trans-splicing. Int.H2A, región intergénica de la histona H2A.

TcRPSEU, Pseudogen ribosomal de *T. cruzi*. E13, elemento repetido

| | | |
|---------|---|-----|
| SIRE | TTAGGA SCTTGTG AAAAGAATGATCGTGGGAGAGCTGGCTAACTTAATTAATGTATGTGTATATCCTGATA- | 72 |
| INT.H2A | .GTA.G.....G.A.G.....-C.A.-.....G | |
| SIRE | AATGAATGCA-TTCTTTAT-GATAC-TTCTACCGTATGAATCTTTTGGGAAGAACGCGACTTTGTAGGGGC | 144 |
| INT.H2A | .C..G.A..T....C..A.G...T.....-.....AA.....AAG..TTA | |
| SIRE | GGGAGCCGATAGAGGCCAGATAATATTTTTAATTTTATTTTGGCATCCACCCACCCCTTGATTCCACCA | 216 |
| INT.H2A |A.....-.....C.....G....C....T....T | |
| SIRE | CGCGGCGGGGTCTTGTGGTTGGAGGACCCCAAAGTCTGCCACTTCGTAAGTAATAATATTTCAAATCCCAAC | 288 |
| INT.H2A |T....T....C.....-.....GG.T. | |
| TCRPSEU |CG.....G.G...T... | |
| SIRE | TGAGGACAAAGGACCATGCTAATGGTCCACAGAATTC-----TATATATT--GAAT | 360 |
| INT.H2A |T....A....C.TG.G.A..A.....A.ATACATATATATATATATA.....TATTCA | |
| E13 |-..... | |
| TCRPSEU | C...TT...A..T..ATTC..C...-----ATA...G...----- | |
| SIRE | GAGAATAAAACATTGAGGCAATTAATAA-----G---AACCTTT TTTCCTAA -CTTGTTGGTTTCCAT | 432 |
| INT.H2A | ..A--G...ACA.....CCTCTGGG---T...C.....T.T.....TT.. | |
| E13 |A.....C----- | |
| TCRPSEU | C...G.....CT...GT....GCCTCTGCT.CCC.T.....C...T-.....CC...TT. | |
| SIRE | AGATAATTT CAGGATCCGGCCAGCTGCCGGGAGGATTATT | 472 |
| INT.H2A |T..... | |
| E13 |A..... | |
| TCRPSEU |T.....-...A.A.-.....C | |

La inserción de SIRE delante de los genes TcP2 β -H1.8 rompe la señal de Trans-Splicing e introduce un nuevo aceptor 3' del SL.

De una biblioteca de ADNc de *T. cruzi* se seleccionaron 10 recombinantes que cubrían toda la región codificante y 5' no codificante de TcP2 β . Entre estos, se encontraron tres tipos diferentes de ADNc que codificaban para distintas variantes de la proteína. En base a las mutaciones puntuales se identificaron como transcritos de H6.4, H1.8 y H1.5. No se aisló ningún transcripto correspondiente a H1.3 (ver capítulo 3) (70).

El análisis de la zona 5' de estos mensajeros mostró que el SL se había unido a los ARNm en posiciones diferentes. En los ARNm transcritos de H6.4 y H1.5 el SL se agregó a 12 bases río arriba del AUG. Por el contrario, en el transcripto derivado de H1.8 se agregaba a 62 bases río arriba del AUG (figura 44 y capítulo 1, descripción de 5D). La comparación de estas secuencias 5' de los mensajeros con su correspondiente genómico determinó el ApG aceptor de SL usado en cada caso. En H6.4 y H1.5, el sitio que actúa como aceptor 3' está precedido 8 bases río arriba por una región de 20 pirimidinas seguidas, incluyendo la secuencia canónica de trans-splicing en tripanosomas TTTCPy y el aceptor ApG (ver figura 29). Por el contrario, en el locus H1.8, esta región fue destruida por la inserción de SIRE en la región del polidT que eliminó la distancia mínima necesaria (4-40 bases) para que el ApG aceptor y la secuencia TTTCPy cumplan su función normalmente en la reacción de trans-splicing (figuras 44 y 45).

Sin embargo, la inserción de SIRE no inhibe la reacción de trans-splicing en H1.8 debido a que dentro del elemento hay una nueva señal de trans-splicing funcional (figura 45). En efecto, 62 bases río arriba del ATG de iniciación y dentro del SIRE se encuentra un nuevo ApG aceptor precedido por 11 bases pirimidínicas incluyendo el consenso TTTCPy. Así en el transcripto H1.8 clonado, el SL es seguido por 38 bases transcritas directamente desde el SIRE. Estos resultados muestran, por primera vez, que una

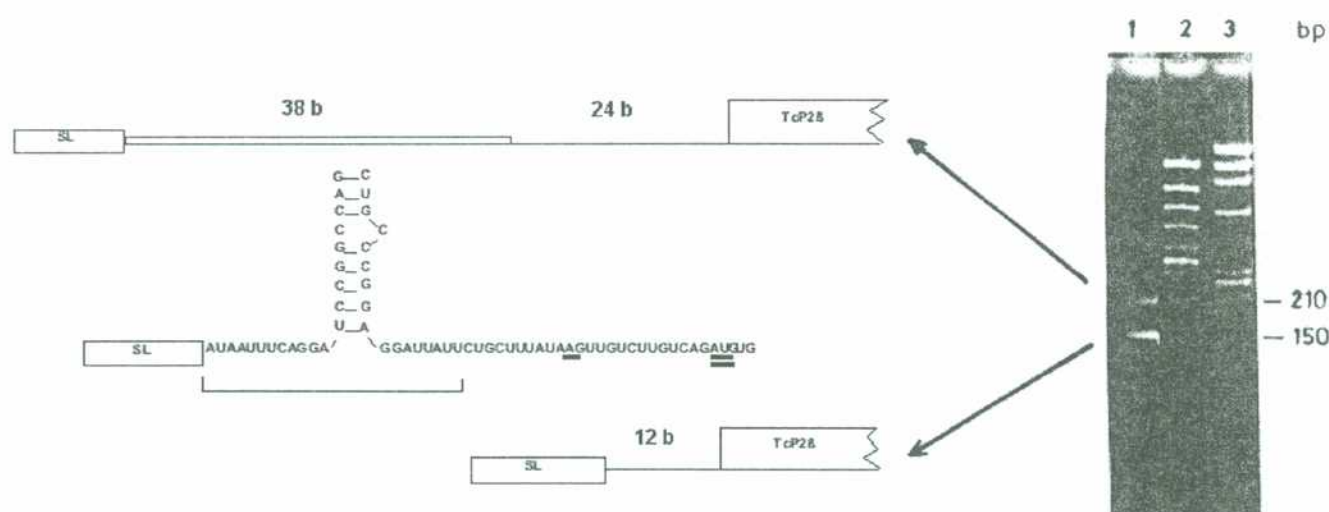
secuencia repetida del genoma de *T. cruzi* lleva secuencias con cierta función en la transcripción de genes (figura 45).

Para confirmar estos resultados se usó un protocolo de ARN-PCR que permitió la amplificación de los extremos 5' de los ARNm de TcP2 β . Se usó un oligonucleótido derivado de la secuencia del SL (ver capítulo 2) y otro complementario derivado de la zona codificante común a todos los mensajeros TcP2 β .

La amplificación dio 2 fragmentos de 166 y 216 pb, indicando que solo se están utilizando los sitios de trans-splicing propuestos (figura 44).

Figura 44

ARN-PCR de los extremos 5' de ARNm TcP2 β . 1, productos de amplificación. 2 y 3 marcadores de PM (ϕ X174 HincII y HaeIII, resp.). Arriba se muestra la representación de l ARNm H1.8 y abajo H6.4. La línea gruesa indica secuencia transcrita de SIRE. Se muestra la



Dentro de las 38 bases transcritas de SIRE, hay una secuencia palindrómica de 18 bases capaz de formar una estructura secundaria en el extremo 5' del mensajero (figura 44). Este tipo de estructuras han sido descritas en policistrones bacterianos y tienen influencia en la estabilidad del ARN (82). Sin embargo, su influencia concreta sobre la estabilidad del transcripto TcP2 β -H1.8 aun queda por ser demostrada. Si nos basamos en los datos de clonado (ver capítulo 3) se puede especular que la transcripción

estructura secundaria de la región 5' del transcrita H1.8

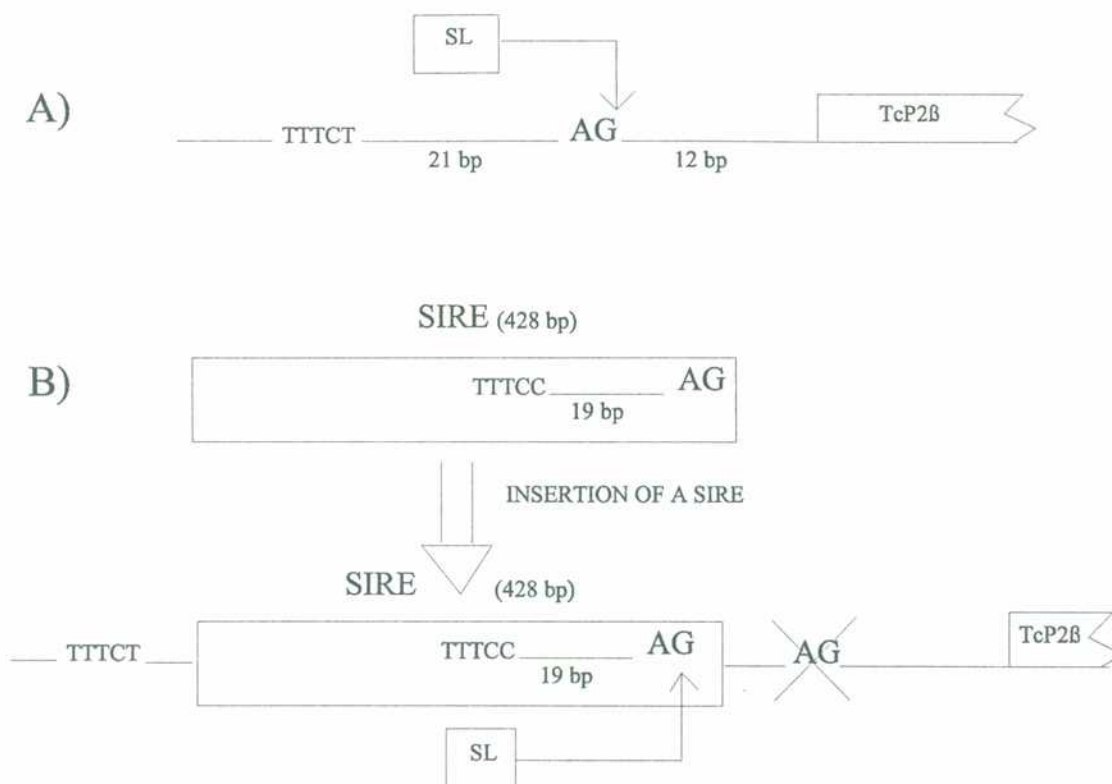


Figura 45-

Diagrama que muestra la interrupción de la señal de trans-splicing en el locus H1.8.

A, Trans-splicing antes de la inserción de SIRE (ej. locus H6.4).

B, Inserción de SIRE, interrupción de la señal original y utilización de la nueva señal ubicada dentro de SIRE.



SIRE: CARACTERÍSTICAS Y ORGANIZACIÓN GENÓMICA.

5.1- Estudio y funcionalidad de SIRE usando como modelo el locus TcP2 β .

Objetivos:
El objetivo de este capítulo es profundizar en el estudio de SIRE, su funcionalidad e interacción con el genoma de T. cruzi.

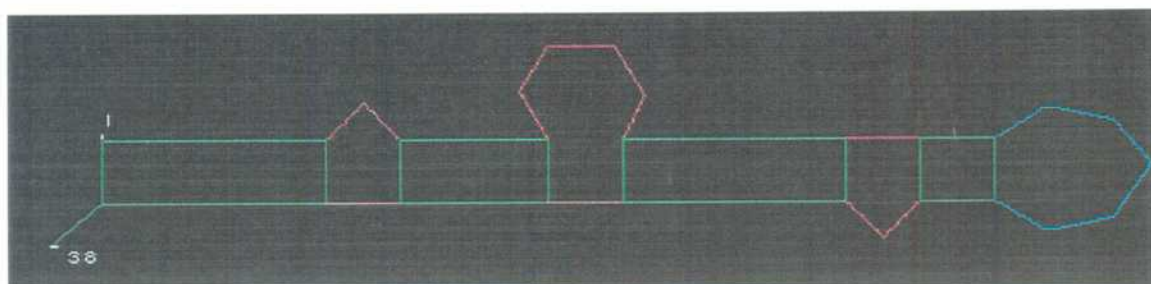
Según se describe en el capítulo anterior, SIRE lleva una secuencia consenso de trans-splicing que es utilizada por el gen TcP2 β -H1.8, dando un mensajero 38 bases más largo que los transcritos desde los otros loci. Además es capaz de formar una estructura secundaria en esa región que podría afectar su estabilidad. Esta estructura secundaria es bastante estable, con una energía libre de -14.6, como se muestra en la figura 44 y 46A.

Esta estructura secundaria se obtiene usando el programa Dnasis. Como se puede observar en la figura 46A se forman distintos loops (marcados en rojo y celeste), que son estructuras conocidas por su sensibilidad al ataque por ARNasas. Este hecho puede ser determinante en la vida media de un ARNm afectando su estabilidad (mayor degradación) y produciendo menos proteína.

Una experiencia de clonado de ADNc de mensajeros TcP2 β mostró que el ARNm H1.8 se encontraba menos representados entre los clones aislados de la biblioteca que el resto de los ARNm expresados desde otros loci (ver capítulo 3). Esto podía implicar que su estabilidad fuera menor de acuerdo a lo propuesto más arriba.

Por esta razón, consideramos al elemento repetitivo SIRE ($\sim 10^4$ copias por genoma) de gran importancia potencial en la expresión genética del parásito,

A



B

```

1  nnnnGAATTC CTCTCTAGGA GAATTTTTTG AGCACAACGA
   cambio A por T en oligo H (HindIII) A
                                     ↑↓
51  TTTAGAGAAT AGTTTAGCTG AAAGGAGTGA TGCTTGCTTG
                                     ↑ OLIGO H
91  GATGAGTGCA AGACCATGGG TGACGTAAAT GCCCTATTCT
   NCO I
131 CTGAAAGCGG ATTGCCTACC TTGCCCTAA CACGAGAGGC
171 GGTTAACGAG TTTCTTCAAA ATATGCAGCG GATTCACTAA
211 GAAACATTTT CACGCACGAA AGCGAAATTA TTATGATTGT
251 TATTATAATA CTTTTTCTTT GTTGTTTTAT CCACTTATTA
291 TTGTTGTGTT AAATTGTTTT TACCTTTTTT CTTTCCAAC
   SIRE SIRE insercion
                                     ↓
331 TTCTTTTATG ATGTCTTTTC TTTTTTTTTT TTTTGCTCTA
   T insercion de T en oligo X (XbaI)
                                     ↓
371 TAAGTTGTCT TGTCAGATGA AGTACCTTGC CGCTTACGCT
                                     ↑ OLIGO X

```

dado que por su dispersión y alto número de copias podría estar influenciando la expresión de un gran número de genes.

Como primer objetivo se planteo entonces estudiar más en profundidad la influencia de este elemento en la expresión de TcP2 β .

Para ello, se está desarrollando un experimento de transfección transiente en tripanosomas que aun no ha concluido. A continuación se detallan algunos aspectos de la estrategia planteada y el armado de los vectores utilizados.

Se han desarrollado variaciones a las técnicas convencionales de transfección para adaptarlas a tripanosomas. Estos métodos incluyen la introducción por electroporación de vectores recombinantes que contienen secuencias de ADN del parásito más algún gen marcador como los que codifican para cloranfenicol acetil transferasa (CAT), higromicina fosfotransferasa (HYG) o neomicina fosfotransferasa (NEO) (83, 84, 85).

Se ha demostrado que no es esencial colocar regiones 3' no codificantes de los genes a estudiar dentro de los vectores de transfección, por dos motivos principales: a) La poliadenilación se encuentra acoplada al trans-splicing, por lo que alcanza con una sola señal y b) la terminación de transcripción ocurre con frecuencia dentro de los vectores plásmidicos utilizados (83, 84, 85).

Dado que nos interesaba estudiar solo como afecta SIRE la expresión del gen, el sistema elegido para trabajar es de transfección transiente. El plásmido utilizado es el **pCAT-Basic** (promega, USA).

El objetivo es colocar la región 5' del gen TcP2 β de un locus sin la inserción con SIRE (H6.4) y de otro locus en donde SIRE se haya insertado (H1.8), para medir el efecto del elemento SIRE sobre el trans-splicing y la estabilidad del ARNm del gen reportador (CAT).

Se diseñaron oligonucleótidos (cebadores) sobre la secuencia del locus H6.4 para amplificar por PCR la región 5' de este locus que carece de SIRE. Estos cebadores amplifican también sobre el templado H1.8 la misma

Figura 46-

A, Estructura secundaria formada en la región 5' del transcripto H1.8 desde la bases 1-38. Generado con el programa DNASIS. Las líneas verdes son bases que aparean y en rojo se ven los loops.

B, Detalle de la secuencia 5' de los genes TcP2 β amplificada por PCR y las mutaciones introducidas para clonar en el plásmido pCAT.

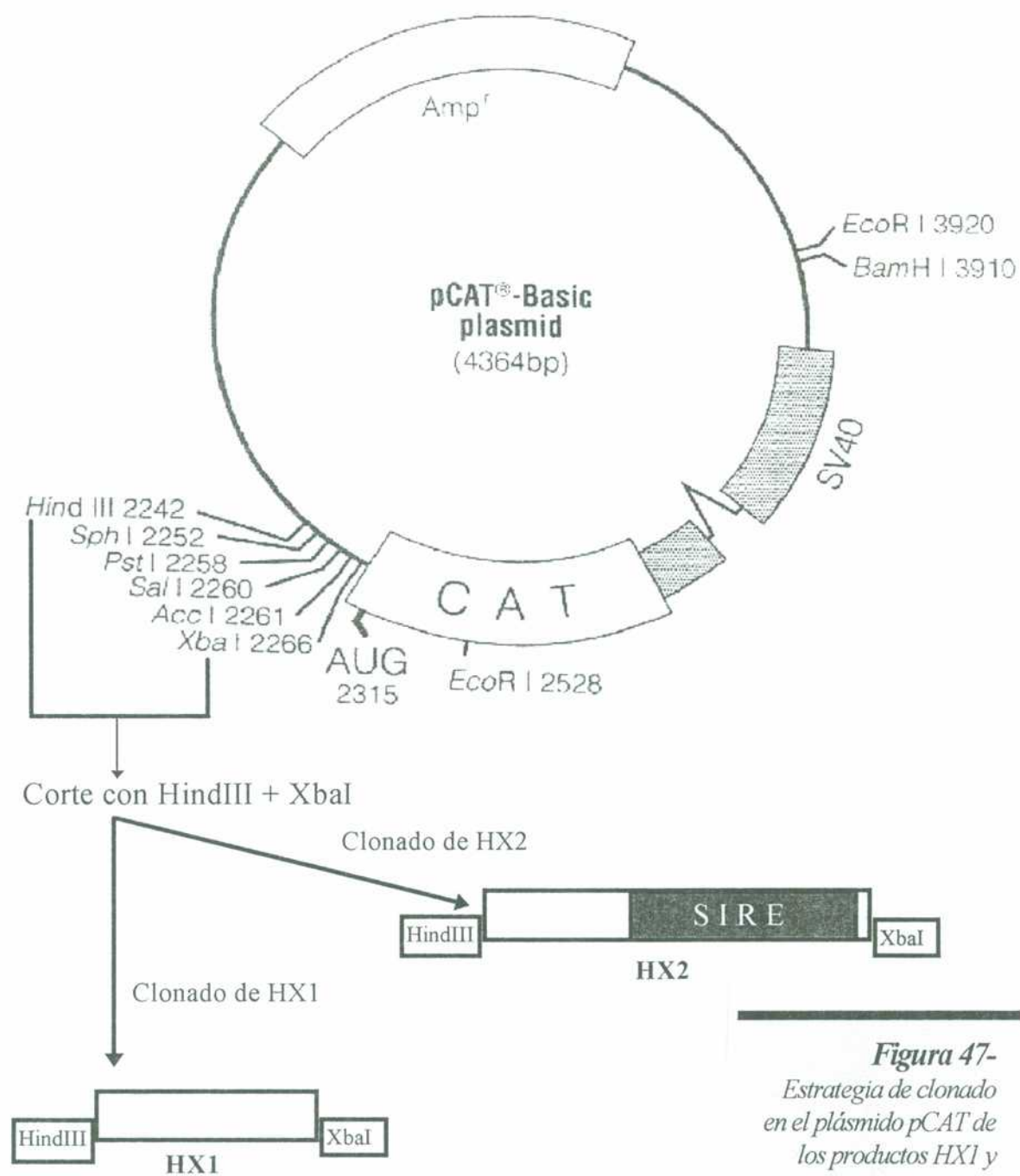


Figura 47-
Estrategia de clonado
en el plásmido pCAT de
los productos HX1 y
HX2, para el estudio de
funcionalidad de la se-
cuencia SIRE.

De esta forma se generaron tres plásmidos para realizar el ensayo:

- 1) pCAT-0 = plásmido control de background, sin productos clonados.
- 2) pCAT-HX1= control basal, medición de expresión del reportador CAT sin influencia de SIRE.
- 3) pCAT-HX2= medición de la influencia de SIRE sobre la expresión del reportador CAT.

En la actualidad, el laboratorio está preparando la infraestructura para llevar a cabo cultivos de parásitos, hecho que va a facilitar los ensayos de transfección.

5.2- Organización y distribución genómica de SIRE en *T.cruzi*: Estructura del elemento y posible función.

La secuencia SIRE se encuentra dispersa y presente en alrededor de 2000 copias en el genoma de *T.cruzi*. Ensayos de PFGE demostraron además que SIRE esta representada en todos los cromosomas del parásito, obsérvese que en la figura 48 se hibridizan todas las bandas cromosomicas separadas. Las diferencias de intensidades pueden estar indicando o bien que hay mas de un cromosoma por banda o que la secuencia SIRE esta representada en muchas mas copias en algunos cromosomas en particular.

Por otra parte, se analizaron distintas cepas de *T.cruzi* y se observaron ligeras diferencias en el patrón de bandas SIRE en los southern blots, tomando como referencia a la cepa Tulahuen. El patrón más marcadamente diferente de bandas se observó en la cepa G (ver

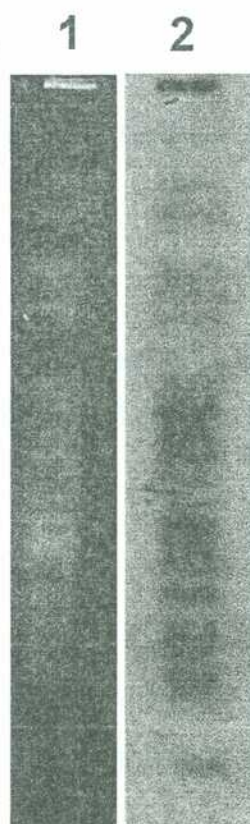
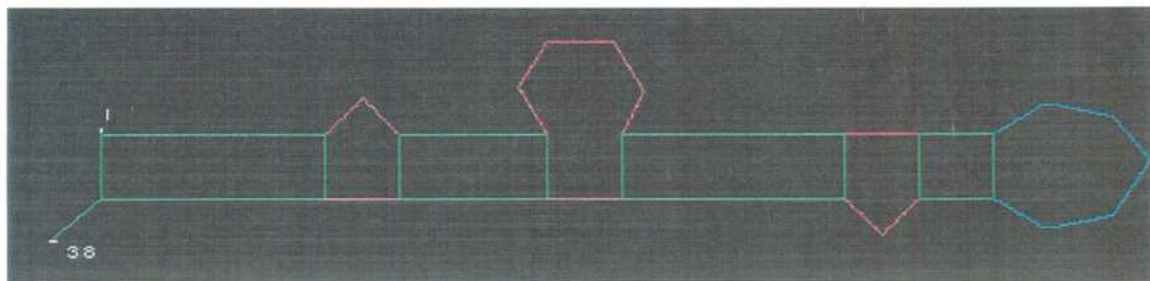


Figura 48-
PFGE de cromosomas
de *T.cruzi*.
1, Gel teñido con Br.Et.
2, Hibridización con
sonda SIRE

A



B

```

1  nnnnGAATTC CTCTCTAGGA GAATTTTTTG AGCACAACGA
   cambio A por T en oligo H (HindIII) A
                                     ↑
51  TTTAGAGAAT AGTTTAGCTG AAAGGAGTGA TGCTTGCTTG
                                     ↑ OLIGO H
91  GATGAGTGCA AGACCCATGGG TGACGTAAAT GCCCTATTCT
   |             NCO I
131 CTGAAAGCGG ATTGCCTACC TTGCCCCTAA CACGAGAGGC
171 GGTTAACGAG TTTCTTCAAA ATATGCAGCG GATTCACTAA
211 GAAACATTTT CACGCACGAA AGCGAAATTA TTATGATTGT
251 TATTATAATA CTTTTTCTTT GTTGTTTTAT CCACTTATTA
291 TTGTTGTGTT AAATTGTTTT TACCTTTTTT CTTTTCCAAC
   SIRE SIRE insercion
                                     ↓
331 TTCTTTTATG ATGTCTTTTC TTTTTTTTTT TTTTGCTCTA
   |             T insercion de T en oligo X (XbaI)
   ↓
371 TAAGTTGTCT TGTCAGATGA AGTACCTTGC CGCTTACGCT
   |             ↑ OLIGO X

```

dado que por su dispersión y alto número de copias podría estar influenciando la expresión de un gran número de genes.

Como primer objetivo se planteo entonces estudiar más en profundidad la influencia de este elemento en la expresión de TcP2 β .

Para ello, se está desarrollando un experimento de transfección transiente en tripanosomas que aun no ha concluido. A continuación se detallan algunos aspectos de la estrategia planteada y el armado de los vectores utilizados.

Se han desarrollado variaciones a las técnicas convencionales de transfección para adaptarlas a tripanosomas. Estos métodos incluyen la introducción por electroporación de vectores recombinantes que contienen secuencias de ADN del parásito más algún gen marcador como los que codifican para cloranfenicol acetil transferasa (CAT), higromicina fosfotransferasa (HYG) o neomicina fosfotransferasa (NEO) (83, 84, 85).

Se ha demostrado que no es esencial colocar regiones 3' no codificantes de los genes a estudiar dentro de los vectores de transfección, por dos motivos principales: a) La poliadenilación se encuentra acoplada al trans-splicing, por lo que alcanza con una sola señal y b) la terminación de transcripción ocurre con frecuencia dentro de los vectores plásmidicos utilizados (83, 84, 85).

Dado que nos interesaba estudiar solo como afecta SIRE la expresión del gen, el sistema elegido para trabajar es de transfección transiente. El plásmido utilizado es el **pCAT-Basic** (promega, USA).

El objetivo es colocar la región 5' del gen TcP2 β de un locus sin la inserción con SIRE (H6.4) y de otro locus en donde SIRE se haya insertado (H1.8), para medir el efecto del elemento SIRE sobre el trans-splicing y la estabilidad del ARNm del gen reportador (CAT).

Se diseñaron oligonucleótidos (cebadores) sobre la secuencia del locus H6.4 para amplificar por PCR la región 5' de este locus que carece de SIRE. Estos cebadores amplifican también sobre el templado H1.8 la misma

Figura 46-
A, Estructura secundaria formada en la región 5' del transcripto H1.8 desde la bases 1-38. Generado con el programa DNASIS. Las líneas verdes son bases que aparecen y en rojo se ven los loops.
B, Detalle de la secuencia 5' de los genes TcP2 β amplificada por PCR y las mutaciones introducidas para clonar en el plásmido pCAT.

región 5' pero con la inserción de SIRE.

Los oligonucleótidos fueron diseñados de forma tal de producir una mutagénesis dirigida durante la amplificación por PCR. El cebador 5', denominado OLIGO H presenta la modificación de una A por una T con respecto al templatado original para introducir un sitio **HindIII** en esa región (figura 46B.). El cebador 3', denominado OLIGO X, inserta una T en el producto amplificado creando de esta forma un sitio de restricción para la enzima **XbaI** (figura 46B.).

Características de los productos amplificados:

1) HX1

Templado: locus H6.4

Tamaño: 310 pb

Descripción: región 5' de TcP2 β **sin SIRE** con la introducción de sitios HindIII en el extremo 5' y XbaI en el extremo 3' del producto amplificado.

2) HX2

Templado: locus H1.8

Tamaño: 740 pb

Descripción: región 5' de TcP2 β **con SIRE** y con la introducción de sitios HindIII en el extremo 5' y XbaI en el extremo 3' del producto amplificado.

En la figura 47 se muestra la estrategia de clonado en el plásmido pCAT. La introducción de sitios artificiales de restricción XbaI y HindIII en el amplificado era necesaria para lograr un clonado direccional en el plásmido.

El clonado fue confirmado por cortes con XbaI + HindIII y con NcoI, que corta tanto dentro de HX1 como de HX2 y dentro del gen CAT, liberando fragmentos de 880 pb si se clonó HX1 y de 1300 si se clonó HX2.

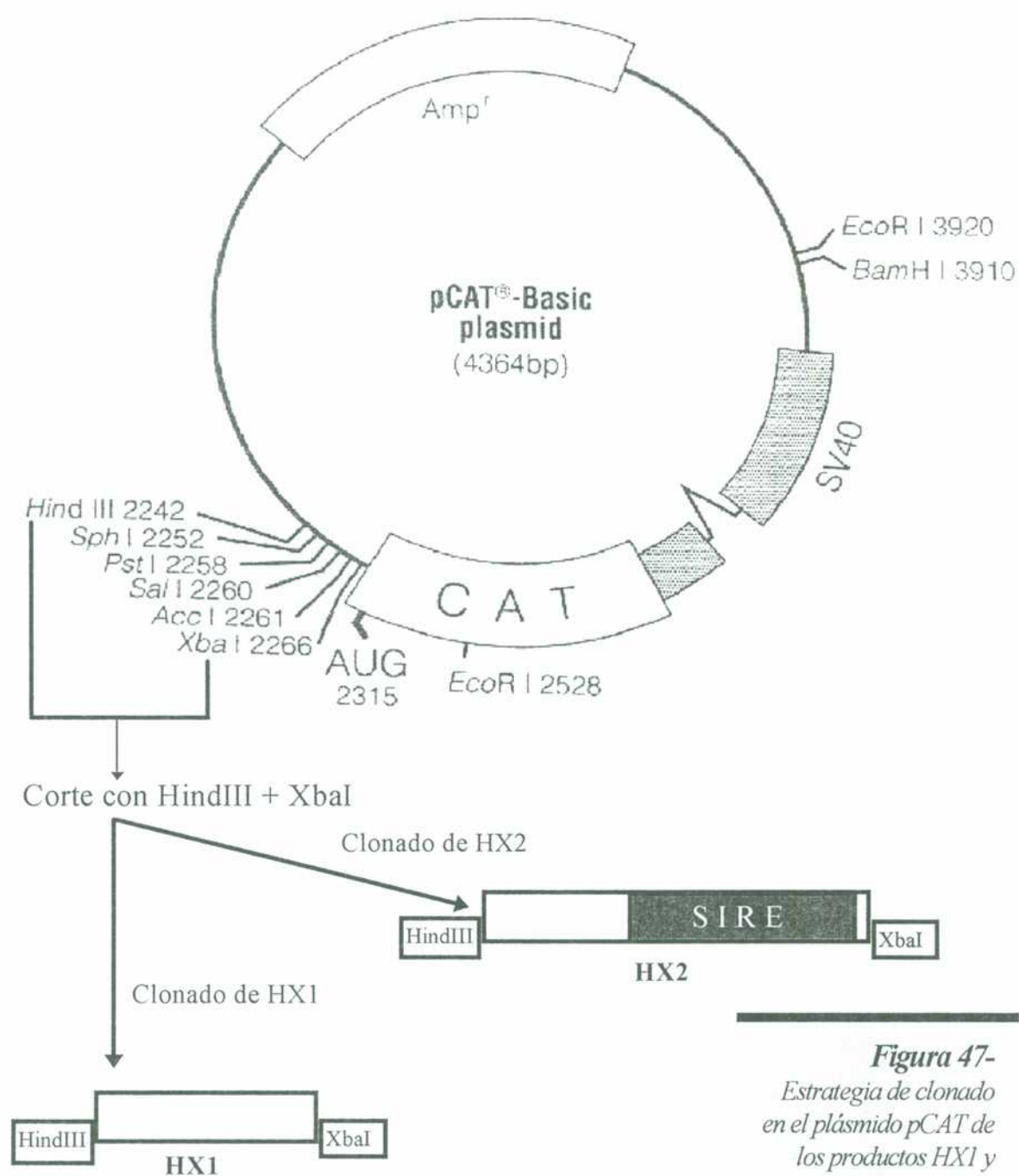


Figura 47-
Estrategia de clonado
en el plásmido pCAT de
los productos HX1 y
HX2, para el estudio de
funcionalidad de la se-
cuencia SIRE.

De esta forma se generaron tres plásmidos para realizar el ensayo:

- 1) pCAT-0 = plásmido control de background, sin productos clonados.
- 2) pCAT-HX1= control basal, medición de expresión del reportador CAT sin influencia de SIRE.
- 3) pCAT-HX2= medición de la influencia de SIRE sobre la expresión del reportador CAT.

En la actualidad, el laboratorio está preparando la infraestructura para llevar a cabo cultivos de parásitos, hecho que va a facilitar los ensayos de transfección.

5.2- Organización y distribución genómica de SIRE en *T.cruzi*: Estructura del elemento y posible función.

La secuencia SIRE se encuentra dispersa y presente en alrededor de 2000 copias en el genoma de *T.cruzi*. Ensayos de PFGE demostraron además que SIRE esta representada en todos los cromosomas del parásito, obsérvese que en la figura 48 se hibridizan todas las bandas cromosomicas separadas. Las diferencias de intensidades pueden estar indicando o bien que hay mas de un cromosoma por banda o que la secuencia SIRE esta representada en muchas mas copias en algunos cromosomas en particular.

Por otra parte, se analizaron distintas cepas de *T.cruzi* y se observaron ligeras diferencias en el patrón de bandas SIRE en los southern blots, tomando como referencia a la cepa Tulahuen. El patrón más marcadamente diferente de bandas se observó en la cepa G (ver

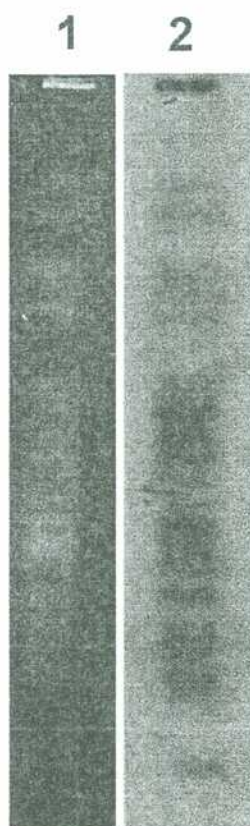
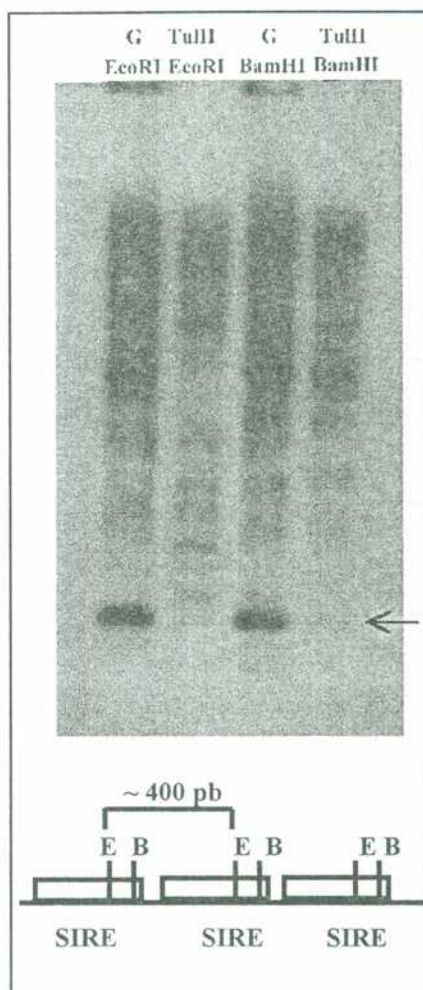


Figura 48-
PFGE de cromosomas
de *T.cruzi*.
1, Gel teñido con Br.Et.
2, Hibridización con
sonda SIRE

figura 49, en donde se compara con el patrón que se obtiene para Tulahuen), cepa que además es la única que muestra una organización diferente de los genes TcP2 β entre las analizadas (ver capítulo 4).

Figura 49-
Southern blot comparativo entre las cepas G y Tulahuen hibridizado con SIRE. La flecha indica una banda de ~400 pb.



Estos datos podrían estar indicando que la secuencia SIRE tiene una organización dinámica dentro del genoma, es decir que efectivamente SIRE tiene la capacidad de moverse por distintos puntos del genoma.

Obsérvese que la flecha en la figura 49 marca una banda intensa de ~400 bp que hibridiza con el corte EcoRI y con el corte BamHI en la cepa G, mientras que en la cepa Tulahuen es tenue. Estas dos enzimas cortan dentro de SIRE separadas por una distancia de 100 bp, de modo que la única explicación posible es que un tandem cola-cabeza de varias copias SIRE se encuentra presente en la cepa G.

Con el objetivo de estudiar más en profundidad la organización y estructura de SIRE en el genoma de *T. cruzi*, se rastreó una biblioteca genómica de la cepa Tulahuen construida en el vector lambda ZAP II.

Se plaquearon 20.000 fagos, luego de hibridización con sonda SIRE, se obtuvieron una gran cantidad de positivos (concordante con el alto número de copias) y se aislaron 50 (figura 50). Finalmente, se seleccionaron 15 fagos recombinantes positivos para trabajar en esta primera etapa denominados SZ.

Los fagos seleccionados se escindieron del fago λ ZAPII, para generar los

plásmidos correspondientes. Estos plásmidos se digirieron usando distintas enzimas de restricción para armar mapas que permitan diseñar la estrategia de secuenciación. A cada plásmido se le realizaron al menos 6 cortes diferentes y la combinación de algunos cortes dobles. Muchas de estas restricciones fueron además hibridizadas con sonda SIRE para

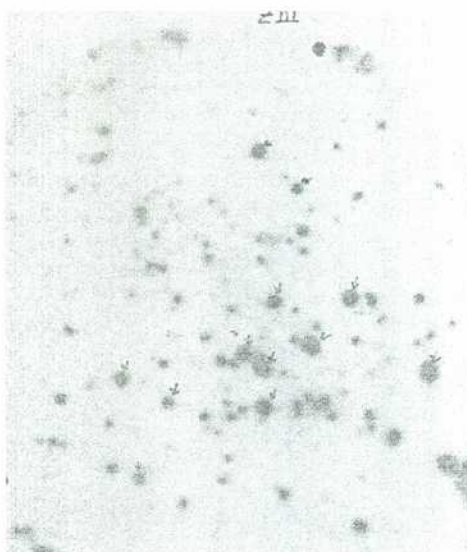


Figura 50-

Rastreo de la biblioteca genómica λ ZAPII de *T. cruzi* con sonda SIRE. Las flechas indican fagos tomados para posterior análisis. Obsérvese que el resultado es coherente con el alto número de copias de SIRE en el genoma.

posicionar este elemento dentro del inserto. El resultado son los mapas que se muestran en la figura 51. En algunos casos, como los clones SZ1 y SZ5 no se pudo posicionar exactamente al SIRE sino dentro de un rango que se señala en la misma figura.

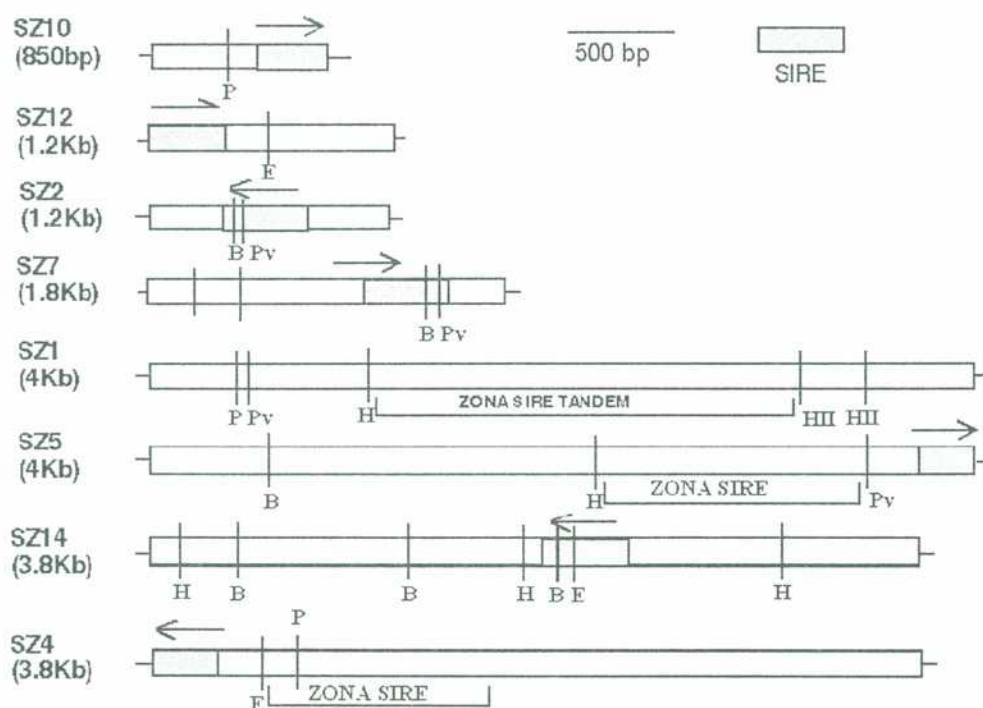


Figura 51-

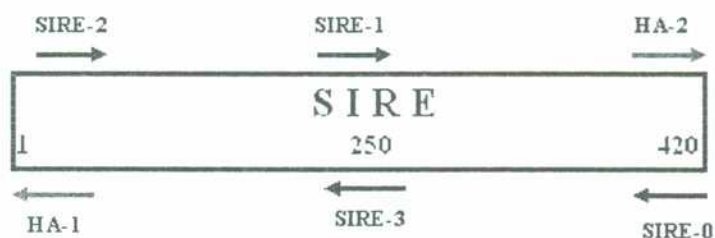
Mapa de algunos de los insertos contenidos en los fagos recombinantes aislados. La s flechas indican la orientación de SIRE en el fragmento genómico. P, PstI. E, EcoRI. B, BamHI. Pv, PvuII. HIII, HindIII.

En la tabla IV, se resumen algunos datos de los fagos recombinantes aislados con los cuales se trabajo.

TABLA IV

| CLONES | MAPA | TAMAÑO | SECUENCIA | Nº DE ACCESO EBI |
|--------|------|--------|-----------|------------------|
| SZ1 | SI | 4000pb | NO | |
| SZ2 | SI | 1200pb | COMPLETA | X 83697 |
| SZ4 | SI | 3800pb | NO | |
| SZ5 | SI | 4000pb | PARCIAL | X 83599 |
| SZ7 | SI | 1800pb | COMPLETA | X 83600 |
| SZ8 | SI | 2500pb | NO | |
| SZ10 | SI | 850pb | COMPLETA | X |
| SZ11 | NO | 5000pb | NO | |
| SZ12 | SI | 1200pb | COMPLETA | X 83698 |
| SZ14 | SI | 3800pb | PARCIAL | |
| SZ15 | SI | 2100pb | PARCIAL | |
| SZ23 | SI | 3500pb | PARCIAL | |
| SZ31 | SI | 900pb | PARCIAL | |
| SZ32 | SI | 2200pb | PARCIAL | |
| SZ36 | SI | 2600pb | PARCIAL | |
| SZ38 | SI | 1300pb | PARCIAL | |
| SZ39 | SI | 2400pb | PARCIAL | |

El mapeo realizado sobre los plásmidos dio evidencias de que en algunos casos SIRE se encontraba en tandem (ej. SZ1, SZ5, SZ4, SZ8). Posteriormente, esto fue confirmado en una experiencia de amplificación por PCR usando como templado los plásmidos SZ. Para las reacciones de PCR y secuenciación se diseñaron varios cebadores sobre la secuencia SIRE como se muestra a continuación:



El PCR realizado con los cebadores SIRE-0 + SIRE-1 mostró que en algunos plásmidos, además de la banda esperada de ~200 pb, se obtenía otra banda más grande de ~800 pb (figura 52). La amplificación sobre el templado del locus TcP2 β -H1.8 (donde el SIRE se encuentra en tandem acompañando a los genes TcP2 β) dio un patrón similar con bandas de 200 pb y 1200 pb, esta última comprende el tandem SIRE que encierra al gen TcP2 β (figura 52). Por lo tanto, los clones SZ1, SZ4 y SZ5 deben contener un tandem SIRE separado por aproximadamente 400 pb de acuerdo al tamaño de la banda mayor.

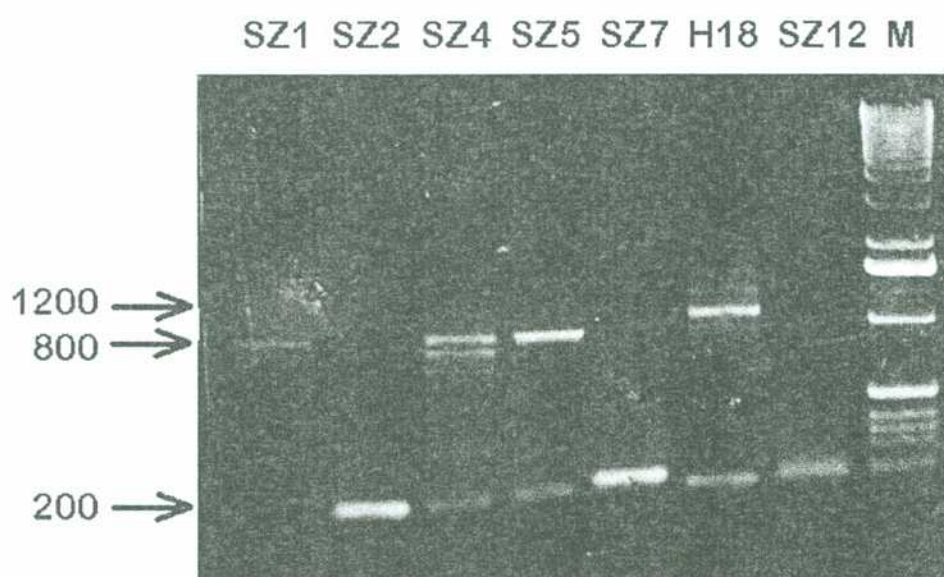


Figura 52-
Amplificación por PCR
de los plásmidos es-
cindidos de λ ZAP con
los cebadores SIRE-0 +
SIRE-1. M, marcador
de PM, 1Kb ladder.

Los insertos contenidos en los plásmidos se secuenciaron utilizando los cebadores SK y T3 (que se aparean dentro del vector, son universales) y los cebadores sintetizados sobre la secuencia del SIRE (mostrados más arriba).

Se obtuvieron las secuencias completas de SZ2, SZ7, SZ10 y SZ12. Además se obtuvo secuencia parcial de SZ5, SZ14, SZ15, SZ23, SZ32, SZ36, SZ38 y SZ39.

Las secuencias completas obtenidas se registraron en el banco de datos de secuencia EBI (European biological informatics). El número de acceso de

estas secuencias se especifica en la tabla IV de resumen.

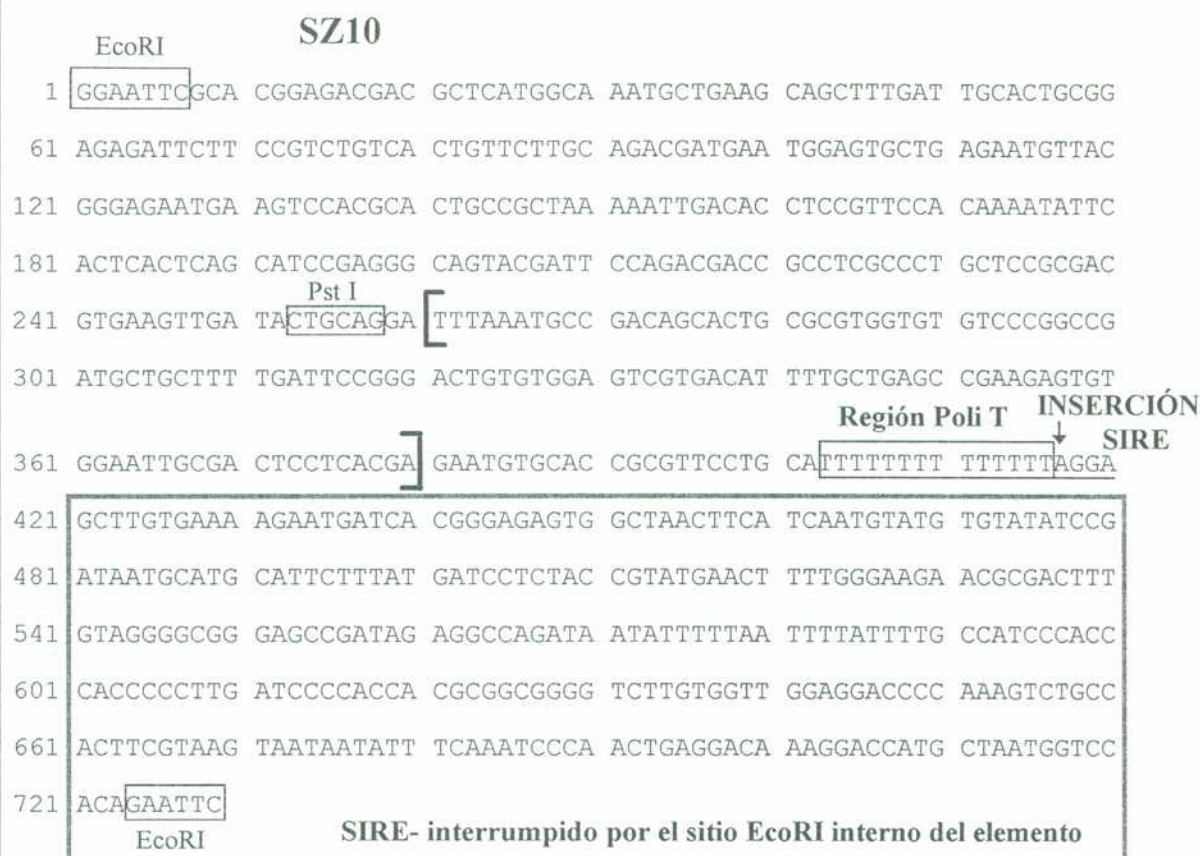
El SIRE presente en el locus TcP2 β -H1.8 se considera como patrón para las comparaciones con las nuevas secuencias obtenidas. De este modo, los SIRE presentes en los SZ secuenciados presentaron la siguiente homología con el SIRE patrón:

- ♦ SIRE SZ10: 95.6 %
- ♦ SIRE SZ2: 91.4 %
- ♦ SIRE SZ7: 69 %
- ♦ SIRE SZ12: 90.7 %
- ♦ SIRE SZ14: 92 %
- ♦ SIRE SZ15: 98.2 % (parcial)
- ♦ SIRE SZ38: 85 %

La secuenciación mostró además que SIRE estaba asociado en todos los casos a loci diferentes, resultado consistente con los datos obtenidos de southern blot, PFGE y número de copias.

Cabe recordar que SIRE se insertaba en el locus H1.8 de TcP2 β en una **región rica en T** y que el elemento en si mismo llevaba una secuencia **AGGA** en sus extremos 5'y 3', quizás duplicada como consecuencia de la inserción. Con este único ejemplo no podíamos generalizar si este evento ocurre siempre de la misma forma y SIRE necesita como target una región de Poli T en el lugar de inserción. El clonado y secuenciación de estos locus SZ permitió clarificar algunos de estos aspectos.

En la figura 53, se muestra detallada la secuencia de SZ10. El recuadro muestra el elemento SIRE insertado en este locus justo después de una región de poli T y seguida por la secuencia AGGA. Desafortunadamente, SIRE se encuentra interrumpido por el sitio EcoRI interno, que también está en el SIRE patrón, y no se puede analizar el extremo 3' de la inserción. El resto de la secuencia adyacente a SIRE no presenta ninguna homología con



el locus TcP2 β -H1.8.

Obsevando las figuras 53 a 59 en conjunto, se puede ver que el elemento SIRE se encuentra siempre inserto en una región de poli T, seguido por la posible secuencia duplicada por la inserción AGGA. En algunos casos se pueden observar variaciones de esta secuencia, pero son cambios G por A. Las transiciones son típicas en las secuencias de *T. cruzi* según se fue observando hasta ahora. En el caso de SZ7, esta secuencia está mutada a AATA, pero es interesante notar que la misma se encuentra tanto en el extremo 5' como en el 3' del SIRE (figura 56).

Estos datos permiten generalizar que este elemento repetitivo utiliza como target para la inserción una región de **Poli T** en el genoma de *T. cruzi*, de la misma forma que habíamos informado para la inserción en el locus TcP2 β .

Esta inserción en regiones de varias T sucesivas hace poco probable la inserción de SIRE en medio de secuencias codificantes, dado que estas

Figura 53-

Detalle de la secuencia SZ10.. Se muestra el sitio de inserción de SIRE en este locus genómico, los sitios de restricción más relevantes y entre corchetes una secuencia (260-380) que presenta homología importante con otras secuencias (ver más adelante)

SZ12

EcoRI

1 GGAATTCAC CGCGCAGTGT TCCGGGAGCG TGTGAGTCCC GACGACTGCT TTGCCCTGCT

61 CTACGACGCG AAATTGACAC [CGCGAGAGTT AACTGCCGAC AGACATGTGC GTGTGCGCGT

121 GTATTGGGTG CTGCTGCTGC TTCTTCTGGG ATTGTGTGGA GTTGTGGTCG GTCTTTTGAG

181 CCAAAGAGTC TGAATTGCG ACTCCTCACG GAGAACATGT] GCCGCGTTCC CTGCGTGTTA

241 ATTTTAAATT ACTTATTTTT GGTTTTGTTG TTTTCCTGGA ATGATAAGCA ACCTTTATTT

301 ATTTAAGGGT TTGGTTGCTG CACACTGTAA TGAGCGAAGT GTTCCCTAT GTTGTGTTTT

361 GTTGCCCGCA CGCTGTCCGG CCTCATTGCT GTGCGTCCCA CGGACTGCAG GAAAGGAGGA

421 TTGGAGGAAC AGGCCGGATG CCGCAATGGA ATGTGAGAGA GTTGTGTGGG CTGGTGGAAC

481 AGGACGATCT TTTTAAATTT TTATTTCTTT TCTTGGTGAA GGTATTATAT GTATACACAA

541 CCACACGATC CTCTTCCATG CTGCATGGCA CAGGAAATAA AAGTGCTCTC TGATAGATTT

601 TTTTAAAAAT TTTCCGGAAT TCTGCTGGAC CCTCAACCCT TTTTTTTTTT TTTTTTTTTT

661 TTCCTGCCTT TTTAAGTTTT GGTTTGTGTG GCGGTGTGCG ACTTTTGCAG TTGTAATGGT

721 TCCGTCTATC TTTTTTGGTG TGCGCTTCCT CTGGCTGCTG CCATTTATTA CAATGCGGTG

781 GCTGTACGTG GTGTAAGTGC TGTCGCTCCC TAC [TTTTTTT TTCTAGGAGC TTTTGAAAAG

841 AATGATCGTG GGAGTGCTGG CTAAGTAAAT TAATGAATGT ACATCCTGAT AAATGAAGGC

901 ATTCTTTATG ATGCTTTCAA CCGTATGAAT CTTTTGGGAA GAACGCGACT TTGTAGGGGC

961 AGGGAGAAGA TAGAGGCCAG ATAATAGGGG TTTTATTTT GCCATCCCAC CCACCCCTT

1021 GAAGCCCACC ACGCGGCCCC CTCTTGTGGT TGGAGGACCT CAAAGTCTGC CACTTCTTAA

1081 GTAATAATAT TTCAGATCCC AACTGAGGAC AAGGGCCATG CTAGTGGTCC ACA [GAATTC

Región de Poli T ↓ **INSERCIÓN SIRE**

SIRE, interrumpido por en el sitio EcoRI interno

Figura 54-

Detalle de la secuencia de SZ12. Se muestra la inserción de SIRE (recuadrado) en la región de polidT. Entre corchetes se señala la región que presenta homología con otras secuencias (ver más adelante, figura 62).

SZ2

1 CACAAACGCA TGGCGCGCAA AAAGAAAAAG AAAAAAAAAAG AAGAAATGAA ATAAGAAGAG
 61 AAAGCGAAGA TCGTATCGAA TTAATTTTAT GGGGATTTTT GTGTGTGTGT GTGTCTGTAT
 121 GTGCGTGGGT TGTGAGCAGG TCGCAGAAAC GGGTTTGTGA TGACCCCTCC TTCCTCTCTA
 181 ATTTTATAAG ACCCCCCTTC GTAAATCTGG CTATTGTTTT CGTTGTCGTT GGGTTTCTAT
 241 GCTTTCACCG CTGCTCTTTT GTTTGTTTGT TTGTTTTTCAT GCACGGAAAG AGGAGGATGA
 301 GATCTGTGTG TGTGTGTGTG TGTTTTTGTT TGGTGGGGTG GGGTGGGGTG GGNTGGGTTT

Región de poli T ↓
 361 TTTTCTAGGA GCTTTTGAGA AGAATGATCG CGGGAGAGCT GGCTAACTTA ATTAATGTAT
 421 GTGTACATCC TGAAAATGAA TGCATTCGGG ATGATACCTT CTACCGTATG AATCTTTTGG
 481 GAAGAACGCG ACTTTGTAGG GGCAGGGAGC CGATAGAGGC CGGATAATAT TTGTTTTATT
 541 TTGCCATCCC ACCCACTCCC TTTTGTTTCC CACCAGGCGG CGGGGTCTTT TGGTTGGAGG
 601 ACCCCAACT CTGCCACTTC GTAAGTAATA ATACGTAAAT CCTAACTGAG GACAAAGGAC
 661 CATGCTAGTG GTCCACAGAA TTTTATATGT TGAATGAAAA TAAACATTG AGGCAATTAA
 721 CTAACGAACC TTTTTCCTT AACTTGTTT GTTCCATAGA TATAACTTCA GGATCCGGC
 781 AGCTGCCCGG AAGATTATTC TCCTTAGAAT CTTATTGTA GCGGAAATAT GTGATGGGAG
 M G

INSERCIÓN SIRE

841 AAGGCAGGCC GGATGACGGA ACACGATTC GGAATTCAG CCTTGTCAC T GACGTTGCCA
 E G R P D D G T R F R N F S L V T D V A

901 CACCATTGTG GAGAAAACCA CCCCACCCA CCCTGAGAGG AGAGAGGAGT CATGCTGGAG
 T P F V R K P P H P T L R G E R S H A G

961 CCTACGCCTT CC
 A Y A F

Figura 55-

Detalle de la secuencia SZ2. Se muestra el sitio de inserción de SIRE (recuadrado en verde) en la región polidT. Dentro de SIRE aparece subrayada y circularizada en rojo, la secuencia la señal de trans-splicing. En azul se marca un posible ATG de inicio de un marco de lectura abierto.

SZ7

EcoRI
 1 GGAATTC CGA TTACCCTCTC AGATGTTAGA GGAATGGGTT TGGGAGCCTG AAATTCTTCA
 61 AAAAATAACA TCTCATTACA AAACGCATGA ATGTTTAACC CAGTAACCTT GTGGAGGCAA
 121 AAGTTAGTAC CCGAAATGAC ATTTATTGGT CGAGATTCTT TAAGACAGTT GGAATTTTCA
 181 CACCTACTCC CTCGAAGTTT TTTCTGCCCC ATTTGCAGGC GGAATGGATC CCAATAAATT
 241 GGATACCGCG GAACTAATGA GGAGTATTCG TTCAAGAATC AGTCCTGACA TTAAATGTGA
 301 GGATGACACA CACTTTGAAT GTGGTTTTGG ACACCTTACT GGTTATGGTG CGGGGTACTA
 361 TGGTTATATG TGGTCAGGTT TTTGCCCTTG ATGTGTACAA TATATCAAGA GTCGTAACGG
 421 CCTTTTGGAT CCTGTTCTTG GGAACCACTA CGTTTCTAAC ATTATTGGCA TTGGGGGCGG
 481 CCAAGATCCC ACTGCCATGC TGAAAAAGTT TTTAGGCAGA GAGCCTAACA GCGATGCTTT
 541 CTTGAAAACCT ATTGGTGTCT ACGATTTAGC TTTTAATAAT GGAAAAAAC GAAAGTAATA
 601 TATAGGGAAA TTACTTTAAG CCTTTTAAAG AATTATTCTG CCACGAAGAG AATGAGGACG
 661 GGTAGGCGGA TTGACCTTTT ACTTGCTGGG GAAAGATTAT AATCGTTATC ATGGGAGTTT
 721 TTGGACTGAA TTTTACGAG GAAAGGCGTC TACCTAAATT AATATATATA TTCTGGCAAG
 781 AAGAAGATAT TTTAAGATTT TTTTTTAAAT CTTATCAATA TGAGCTTGCG AAAGAGAATG
 841 ATCGCAGGGA AGAGCTGGCT AACTTATCCA TAAATCTTGA TAAACGAATA TACTTCTTTA
 901 CAACGCTTTT CTACCGTATG ATTTTTTGGG AAAAACACGA CTTTGTAGGG GCAGGGAACC
 961 GATAGAGGCC AGATAAGATT TCTTTTTATT TTGCCATCCC ATTCACCCCC TTGATTCCCA
 1021 CCACGCGGCG GGGTCTTGTG GTTGGAGGAC CCCAAAGTCT GCCACTTCTT AATTAATAAT
 1081 ATATATATAT ATATATATCT CCAGCTCGAA TAGAGTTGCT GCCATGGCTG CGATATTGTC
 1141 CATCTACCAT GCATGTTTTG GGAGCAAATA TCGAGGAAGT GACGAACTGA CTACCTTTCC
 1201 ATTCACCTTT TTCTAATTCT CTTCTTGTTT TATTTGTATA GATAGTTTGA GGATCCGTCC
 1261 ACAGCTGGG GCAATAATTT TTAGATTTTT TTTTGACCA TTTTCACTC ACATTACTCG
 1321 CGCTATAACG TTTTCTTAA TTTCTATTCC ATGTTTTGTT GGTTTTAGTC TCTCTATTCT
 1381 CGTCTGTTTC TTTATATATA TCTAAACGGG GAGGAGATTT TGATACCAAT GGATTGGTCT
 1441 TCTACGTCTT CTGATGAGGA GCTCCGAAGC ACTCTGCGCA CAACGCCAGA TATAGCAGGT
 1501 CGAGGGAGCG TCGTACCAAC TT

Región de Poli T
 INSERCIÓN SIRE
 INSERCIÓN SIRE
 PvuII
 BamHI

Figura 56- Detalle de la secuencia de SZ7. Dentro de SIRE (recuadra verde) se subraya en negro una región (TA)_n repetida., luego de la cual hay una zona que pierde homología con SIRE (punteado). Finalmente la homología se retoma en la zona que contiene las señales de trans-splicing (subrayado y circularizado en rojo).

SZ15

```

1 TATCGCTTTC GCATCTGCAT CCACCATGAC GTTGGTGACA ACACTTCGCT GCAAGGAGCG
61 GCGAATAAAG GCAGCTCGAA ATCGCACGCC ATGACGTGGG AGCTGCGACG GATATATGAG
121 TTTTGGGACT CTCGCGGAAT ACAGGCAACT TTTGCCACG TGCGGTTCG ACAAAACCCC
181 GCAGACGGCA TATCACGCGG TCGTGTGTTT ACACTTCAGG ACTTGGCGAA GGGGTGGAAC
241 TTGCGAAGGG GTCGCGCGGG GTCTTGTGGT TGGAGGACCC CAAAGTCTGC CACTTCGTAA
301 GTAATAATAT TTCAAATCCT AACTGAGGAC AAAGGACCAT GCTAATGGTC CACAGGGAAT
361 TCC

```

Figura 57-

Detalle de la secuencia SZ15. Se recuadrada la región parcial de SIRE que está asociado al pseudogen ribosomal (ver texto y cita 80). En rojo aparece un sitio PstI

SZ38

```

1 AGCACTGATG TATACTATTA GACTGTGACA CAGTTTCTGA GACATGCCTG TGCTGTGCCT
61 GCACGTCTGG ATCTTTTCT GTTGGACGAA CATTATCTGT ACCGCACCAC CCTCGCGATC
121 CTTGTGTATC GGTGTTGCT GTACACAGG CATGCACTGA GTTCTCTTCC ACGGTGCTAT
181 TGCTCGCAAG AAGATGTGTT GCAAACCTCTA GACCCATCTT CTTTAGACTT TCTACATGTC
241 TCCGTCACIT TTCTGTCGTA TTTTAGCGAA TGATAAATTG TGTGGTATAG CATGCGTTCT
301 ATTCGCTCCA TGCGAAGCGT GTGGCAAGTA ATCGTCCCTC TCATTAGTAG TATGACGCAA
361 CAACACGTGG CTTTATTTG GGTCTTGTG AAAAGAATGT ACGCGTGGTA GCTGGTAGCT

421 GGCTAACTTA ATTAATGTAT GTGTATATCC TGATAAATGA AGGCATTCTT CATGATACTT
481 TCTACCGTAT GAATCGGCTG GGAAGAATGC GACTTTGTAG GGGAGAGCAG GGAGCCGATA
541 GAGGCCGGAT AATATTTGTT TTTATTTTGC CATCCCACCC TAACCCCTT TTGATTCCTA
601 CCACGCGGCG GGGTCTTGTG GTTGGAGGAC CCCAAAGTCT TGCCACTTCG TAAGTAACAA
661 ATACTTCAAA TCCTAACTGA GAGGGNGAT GCTCGGAATT CC

```

Figura 58-

Detalle de la secuencia SZ38. Se recuadrada SIRE y su sitio de inserción.

SZ31

A

```

1 AANAANTANT TNNTNCANTA NANAACNAAN ANNGNANNAA TCNCANAACN TNAANTAACC
61 NATTNACNAA TNANANAANN AGANNANNAN TTTCCCCAGA ANANNAACA CNAANANNAC
121 CNANTNANAN AAAGANNANA NNAANNNNAN CNANCTANAN CTNAAAATNA AAAATAATAC
181 ANCANCNNAT AANNACTCAN TNTACNNNAT ACNCCTNNCC TGNATTNNCG GNCNCTTANT
241 ANTTTTTACC NANCANANNG NNACTATCCN CCNNTANGGN AGANATNNAA AGCANCNGNC
301 NCGGNNACNC CTNAAAANCA AGTTTTAAAG GAATANTTTA TNATGNNGNC ANGNNGAANA
361 ANNATCCGNT CCTTNTANAT GGAATTGTNA CAGGNCATA TGNTCTACNA NCGTNCCCAT
421 TTNTTACNGT AACACTTTGC CGNACAACAT GATAACGAAA CCTAACTTCN CGNGAGCTTT

481 TATTAGGAGC TTGTGAAANG AATGATCCCG GGAGAAGCTG GCCAACTTAA NTAATGTNTG

541 TGTATANCCT GATAAATGAA TGCATTCTTT ATGATACTTT CTACCGTATG AATCTTTTGT
601 GGAAGAAGCG CGACTTTGTA GGGGCGGGAA CCGATAGAGG CCGGATAANT TTTAATTTTA
661 TTTTGCCTTC CCACCCACCC CCTTNTGATT CCCACCTGCG GGGGGGGTCT TGTGGTTGGA
721 GGAGCCCCAA GTCTGCCACA TCGTAAGTAA TAATAATCAG ANTCTAANTG AGGAAAAAGG
781 ACCATGCTAA TGGTCAAAAG AAATCGATAT CAAGCTTATC GATACCGTCG ACCTCGAGGG
841 GGGGCCCGGT ACACAANTCG CCTATAGTG ANCNNGAANN N

```

B

```

1 GGAATTCAC CGCGCAGTAT TCCGGGAGTG GTGTTAGTCC GACGACTGCT TTGCCCTGCG
61 CACGACGCGA AATTGACACT GCGCGAGTTA ACTGTTGACG CACTGTGCGT GTGTGCGTGT
121 ATTGGGTGTT GCTACTTCTT CTGGGATTGT GTGGAGTTGT GGCAGACTTC TGAGCCAAAG
181 AGTGTGGAAT TCGCATCCTT CGCATGGAAC ATGTGCCGCG TTCTCTGCTT GTTCTCTTTA
241 TTATTTACAT TTACTGTTGG TTTGTTGTTT CCTGGAATGA TAATTATCCT TGCCTTCCTT
301 ATTATTTAAT GCGT

```

Figura 59-

Detalle de la secuencia SZ31. A., Región 5' del inserto recombinante. Se recuadrada SIRE y su sitio de inserción. B., región 3' del inserto recombinante. Faltan bases para completar el overlap con la secuencia A. Esta zona presenta homología con SZ12 (ver figura 60).

zonas repetitivas de T intercaladas con G o A son típicas de regiones intergénicas no codificantes en *Trypanosoma*. Las regiones codificantes tienen porcentajes de G+C cercanos al 50 %, pero en las regiones intergénicas donde predominan las repeticiones de T, el porcentaje de G+C puede caer a valores tan bajos como de 20 % (ver capítulo 4, figura 35). Este hecho es importante porque SIRE no produciría inactivaciones insertionales de genes, lo cual realza la hipótesis de que podría estar cumpliendo alguna función importante en el genoma. En esta dirección, hay que apuntar también la alta conservación que existe de la secuencia en si misma (valores de 90 % de homología respecto al SIRE patrón), la única excepción estudiada es en SZ7 (ver más adelante).

En SZ2 podría haber otro ejemplo del mismo fenómeno que ocurre en el locus TcP2 β -H1.8. Como se puede ver en la figura 55, en la región 3' del SIRE se encuentra la región de polipirimidinas y el aceptor AG que son las señales necesarias para que ocurra trans-splicing, río abajo del SIRE se puede encontrar un marco abierto de lectura que llega a 46 aminoácidos de longitud (recuadrado en azul se muestra el ATG de inicio) interrumpido por la finalización del inserto clonado. Es probable que si este marco se expresa, use las señales que se encuentran dentro del SIRE, dado que la secuencia entre el SIRE y el probable ATG de inicio no contiene otras señales de trans-splicing. Si esto ocurre, sería un caso similar al descrito para TcP2 β -H1.8.

En SZ7 también se encuentran conservadas las señales pero no hay suficiente secuencia río abajo de SIRE como para determinar si hay algún marco abierto de lectura, aunque en esa zona también hay otras posibles señales.

El elemento presente en este SZ es el que presenta la menor homología con respecto al SIRE patrón. Esto se debe a una zona restringida que no mantiene la homología. En la figura 56, se marca una región con línea negra de dinucleótidos TA repetida, a partir de allí la zona marcada en línea

punteada no conserva la homología. Esta homología se retoma solo a final de la región 3' (zona que justamente es la que se transcribe y contiene la señal de trans-splicing en el caso del locus TcP2 β). Es probable que exista alguna presión de selección para mantener la secuencia de esta zona de SIRE.

La región que contiene (TA)_n repetido ha sido descripta previamente como una secuencia favorable a producir recombinaciones tanto en *T.brucei* como en *T.cruzi* (J. Requena, comunicación personal). Se puede especular que el hecho de que este dinucleótido TA se encuentre más repetido (10 veces) en SZ7 que en el SIRE patrón (3 veces) indique que ha ocurrido algún tipo de recombinación mediada por esa zona y por eso no se mantiene la homología río abajo.

En SZ15 (figura 57) esta contenida una porción 5' de SIRE, que habíamos detectado previamente en una búsqueda en el banco de datos EBI, asociada al pseudogen de ARN ribosomal 24S α (ver TCRPSEU en figura 43, capítulo 4). SZ15 extiende la secuencia descripta por Vieira de Arruda et al. (80) hacia la región 3' de ese locus genómico y la única secuencia en común entre la descripta y SZ15 esta dada por esa región parcial de SIRE y 80 bases más hasta el sitio PstI (ver figura 57). Como se puede observar, SIRE se conserva exactamente hasta una zona rica en G+C (posición 260 en la figura 43), a partir de la cual la homología se pierde completamente. Esta región de SIRE acompaña a la zona del ARN ribosomal que se transloca fuera del locus que contiene los ribosomales generando un pseudogen. No se puede afirmar con certeza si este SIRE estuvo directamente involucrado en la translocación y generación del pseudogen, pero todo el hecho es sugestivo.

Por otra parte, cuando se analiza el resultado de hibridización en southern blot de ADN genómico con una sonda de SZ15 que no contiene la región de SIRE (es decir la secuencia hacia abajo de SIRE en el locus), se pueden observar al menos 5 sitios genómicos en donde la secuencia está presente .

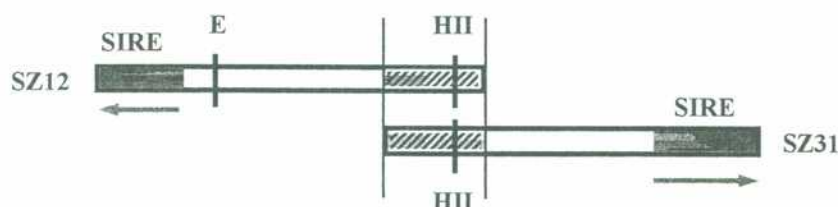
Los resultados reportados por Vieira de Arruda et al. con una sonda ubicada hacia arriba de SIRE en el locus muestra que esa secuencia esta contenida en esa única región genómica (80). Tomados en conjuntos los datos pueden estar mostrando al SIRE presente en ese locus como una especie de “bisagra” que actúe permitiendo recombinación entre distintos puntos del genoma.

Cuando los distintos SZ se comparan entre si se puede observar un grado de homología entre SZ12 y SZ31 (una región entre las posiciones 1 y 280 tomando como referencia a SZ12) de 83 % sin tener en cuenta a SIRE (ver figura 60). SZ12 se extiende 900 bases hacia arriba de esa región y SZ31 se extiende 600 bases hacia abajo de esa región. Es probable que ambos sean loci homólogos duplicados con SIRE insertado en distintas posiciones.

Figura 60-

Comparación entre SZ12 y SZ31 (figura 59, parte B). Obsérvese que la zona que contiene SIRE en SZ31 está en el otro extremo del inserto recombinante (figura 59, parte A; hacia 5' de la secuencia comparada) y en SZ12 se encuentra a continuación (hacia 3') de la secuencia comparada. Esto se esquematiza en la parte inferior: La zona sombreada en ambos SZ corresponde a la secuencia comparada. Las flechas indican la orientación de SIRE. E, EcoRI. HII, HincII.

| NO. | TARGET FILE | DEFINITION | Match% | Over. | INIT | OPT | | |
|--|-------------|--|--------|------------------|------|-----|-----|-----|
| 1 | SZ12.SEQ | | 83.5 | 310 | 240 | 826 | | |
| 83.5% identity in 310 bp overlap, initial score: | | | 240 | optimized score: | | | | |
| 826 | | | | | | | | |
| | | | 10 | 20 | 30 | 40 | 50 | |
| SZ31.SEQ | | GGAATTCCACCGCGCAGTATTCGGGAGTGGTGTAGT-CCGACGACTGCTTTGCCCTGC | | | | | | |
| SZ12.SEQ | | GGAAATTCACCGCGCAGTGTTCGGGAG-CGTGTGAGTCCCGACGACTGCTTTGCCCTGC | | | | | | |
| | | | 60 | 70 | 80 | 90 | 100 | 110 |
| SZ31.SEQ | | GC-ACGACGCGAAATTGACACTGCGCGAGTTAACTGTTGACGCAC-TGTGCGTGTGTGCG | | | | | | |
| SZ12.SEQ | | TCTACGACGCGAAATTGACACCGCGAGAGTTAACTGCCGACAGACATGTGCGTGTGCGCG | | | | | | |
| | | | 120 | 130 | 140 | 150 | 160 | 170 |
| SZ31.SEQ | | TGTATTGGG---TGTTGCTACTTCTTCTGGGATTGTGTGGAGTTGTGG-CAGACTTCTGA | | | | | | |
| SZ12.SEQ | | TGTATTGGGTGCTGCTGCTTCTTCTGGGATTGTGTGGAGTTGTGGTCTGGTCTTTTGA | | | | | | |
| | | | 180 | 190 | 200 | 210 | 220 | 230 |
| SZ31.SEQ | | GCCAAAGAGTGTGGAATTGCGACTCCTCGCATGGAACATGTGCCGCTTCTCTGCTTGT | | | | | | |
| SZ12.SEQ | | GCCAAAGAGTCTGGAATTGCGACTCCTCACGGAGAACATGTGCCGCTTCCCTGCGTGT | | | | | | |
| | | | 240 | 250 | 260 | 270 | 280 | 290 |
| SZ31.SEQ | | CTCTTTATTATTTACATTTACTGTTGG-TTTGTTG-TTTCCTGGAATGATAATTATCCTT | | | | | | |
| SZ12.SEQ | | --AATTTTAAATTAC--TTATTTTGGTTTGTGTTTTCCTGGAATGATAAGCAACCTT | | | | | | |
| | | | 300 | 310 | | | | |
| SZ31.SEQ | | GCGTTCCTTATTATTTAATGCGT | | | | | | |
| SZ12.SEQ | | TATTTATTTAAGGGTTTGGTTGC | | | | | | |



El mismo caso se puede observar entre SZ14 y SZ23. En este caso hay un 70% de homología entre ambos en las posiciones 80 a 400 tomando como referencia a SZ23 (figura 61).

```

NO. TARGET FILE DEFINITION Match% Over. INIT OPT
1 SZ14HA1.SEQ < 69.6 326 478 656
69.6% identity in 326 bp overlap, initial score: 478 optimized score:
656
COMPLEMENTARY key sequence

          90      100      110      120      130
SZ23.SEQ   CCCTCTTCGTGAGCTCGTCTCTCTGTCATCGTCGAC-GATATTGCTGCACG-TACTGGTGC
          :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
SZ14HA1.SEQ TCCACACTGGAAACTCTTTGTTGATGTCATCGGCTGATACTCTACTCGATACTGGTGT
          10      20      30      40      50      60
SZ23.SEQ   ACCTCCTCCTCCAAGAACGCCGCAATCTCTCGGTTTTCCTTCTTGCTCACATTATCGGAA
          :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
SZ14HA1.SEQ ACCTTTTCTTCCAAACTCCGCAATTTCT---TTTTCAGCGTCGGTTTCATTATTGGGA
          70      80      90      100      110
SZ23.SEQ   TTTACAACATCACATCTCTGCCAGTCATTTCATCGGCGTGTCTGCGTGTGCACGTAA
          :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
SZ14HA1.SEQ TTGGCGGGACCATCTCTGCCAGCCAGTTAGCGGCTTGTGCTTTGTGCTGCACATAA
          120     130     140     150     160     170
SZ23.SEQ   ATAATCTTCCACGACATGTCTCGAGATAATTTCTTCCAGCCATTAAATATGCCGCCAGG
          :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
SZ14HA1.SEQ ACAATCTCCACGACATGTCTCGGATAAGGCCTCCCAATCATTAATATGCCGCCAGC
          180     190     200     210     220     230
SZ23.SEQ   CACTCAGTGAAGTGCTCACAGTGTGGCGGTGGTGTGGTGGCCAC-CCGCCGTAGCCAT
          :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
SZ14HA1.SEQ TGTTCAGTGAACAGTTCACAGTACTGGTTAGGGTGTGGTGTGCTCACTTAGCCGTAGTGAT
          240     250     260     270     280     290
SZ23.SEQ   CCGCAGCCCAACCAGCGTCACTGGATT
          380     390     400
SZ14HA1.SEQ CTGCAG-CCAACCATTATCTTCGGGTT
          300     310     320

```

Figura 61-
Comparación de se-
cuencias entre SZ14 y
SZ23

Los restantes SZ no comparten homología entre si, salvo por una corta zona presente en los clones SZ10 y SZ12. Esta zona se marca entre corchetes en las figuras 53 y 54. Este resultado se obtuvo haciendo análisis de HARR PLOT usando el programa de análisis de secuencias Dnasis. El HARR PLOT es un gráfico que se realiza sobre dos ejes, en el eje X se coloca una secuencia analizar y en el eje Y se coloca la otra. El programa toma una ventana definible de n bases que va corriendo a lo largo de las secuencias y coloca un punto en el gráfico cuando de esas n bases de la ventana un porcentaje también definible presenta homología. Este método permitió identificar una zona de homología entre SZ10 y SZ12 diferente a la zona de homología definida por los SIRE contenidos en ambos clones. (figura 62 A.)

B)

CUTOFF : 100 KTUP : 4
 TARGET : GenBank database RELEASE : R74.0
 GROUP NAME : Invertebrate
 NO. TARGET FILE DEFINITION Match% Over. INIT OPT
 TRBFL160 TRYPANOSOMA CRUZI FL-160 EPITOPE GENE, 65.6 122 138 206
 65.6% identity in 122 bp overlap, initial score: 138 optimized score: 206

```

      10      20      30      40      50
SZCOMM.SEQ AGGATTTAAATGCCGACAGCACTGCGC---GTGGTGTGTCCCGCCGATGCTGCTTTTGA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TRBFL160    TGGCTCTAATTGGTGACAGCACCGTGCATGGGTGTGTGCTCGGGTGTGTTGCTGCTGC
      1020    1030    1040    1050    1060    1070
      60      70      80      90      100     110
SZCOMM.SEQ  TTCCGGGACTGTGTGGAGTCGTGACATTTTGCTGAGCCG----AAGAGTGTGGAATTGCG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TRBFL160    TTCTGGGGCTGTGGGGCACCGGCTCTCTGCTGAGGCGTGAGGAAGGAGTGTAAATTGCG
      1080    1090    1100    1110    1120    1130
      120
SZCOMM.SEQ  ACTCCTCACGAGAAT
      : : : : :
TRBFL160    GCTTTTGTG-GAAC
      1140

```

C)

CUTOFF : 100 KTUP : 4
 TARGET : GenBank database RELEASE : R74.0
 GROUP NAME : Invertebrate
 NO. TARGET FILE DEFINITION Match% Over. INIT OPT
 TRBGP85A TRYPANOSOMA CRUZI GP85/SIALIDASE HOMOL 61.9 134 124 182
 61.9% identity in 134 bp overlap, initial score: 124 optimized score: 182

```

      10      20      30      40      50
SZCOMM.SEQ AGGATTTAAATGCCGACAGCACTGCGCGT-GGT--GTGTCCCGGCCGATGCTGCTTTTGA
      : : : : : : : : : : : : : : ^ V : : : : : : : : : : : :
TRBGP85A    TGTATCTAATTGGTGACAGCACCGTGCATGGGTATGTGTCTCGAGTGTGTTGCTGCTGC
      420      430      440      450      460      470
      60      70      80      90      100     110
SZCOMM.SEQ  TTCCGGGACTGTGTGGAGTCGTGACATTTTGC---TGAGCCGAAGAGTGTGGAATTGCG
      : : : : : : : : : : : : : : V : : ^ : : : : : : : : : :
TRBGP85A    TTCTGGGGCTGTGGGGCATTGCGGCTCTCTGCTGAATGAATTGGGGAGAGAAAAATTGCG
      480      490      500      510      520      530
      120
SZCOMM.SEQ  -ACTCCTCACGAGAAT
      : : : : :
TRBGP85A    AAATTCTGTGGAGCAA
      540      550

```

Esta zona de 130 pb es 76.2 % homologa entre los dos clones. Cuando esta región se toma por separado para realizar una comparación de secuencias en el banco de datos del EBI contra otras secuencias de *Trypanosoma* se obtiene los siguientes resultados (ver figura 62 B y C.):

- * 65.6 % de homología con una región presente en el gen del Antígeno FL-160 de *T. cruzi*
- * 62 % de homología con una región presente en el gen de la proteína de superficie GP85/Sialidase.

Esta secuencia está denominada como SZCOMM en la figura 62 A. La región se encuentra a distinta distancia con respecto al SIRE en ambos SZ y no se encontró SIRE en las zonas circundantes de FL160 y GP85/Sialidase, por lo tanto es improbable que esta secuencia guarde alguna relación con el SIRE. Puede ser, sin embargo, que esta secuencia sea un pequeño elemento repetitivo disperso por el genoma como lo es SIRE. Una hibridización en un Southern blot genómico de *T. cruzi* con la banda de 600 pb EcoRI de SZ12 que contiene esta secuencia parece apoyar esta idea (ver figura 64 A., calle ...).

Los datos provenientes de los SZ muestran un panorama más claro en la organización del SIRE, sin embargo queda aun por resolver que sucede en los SZ en los cuales SIRE se encuentra en tandem. Esto puede dar la clave para entender si SIRE forma parte de un elemento todavía mayor del tipo de retrotransposones además de moverse como un elemento autónomo.

Resumiendo los datos obtenidos podemos decir:

- 1) SIRE se encontró asociado a loci diferentes en todos los casos y distribuidos en distintos cromosomas.
- 2) Los elementos SIRE conservan una gran homología de secuencia entre si.
- 3) El Target de inserción de SIRE es siempre una región de Poli T
- 4) SIRE parece no interrumpir secuencias codificantes.

Teniendo en cuenta estos resultados se puede proponer la siguiente estructura de SIRE y un posible esquema de inserción en el genoma (figura 63 A y B.).

Básicamente el SIRE se puede dividir en tres regiones:

- ♦ Una región 3' que contiene las 38 bases que se transcriben y la secuencia adyacente de polipirimidinas que actúa como señal de transplicing. Esta región se conserva aun cuando otras zonas de SIRE no lo hacen (ver SZ7, figura 56).
- ♦ Una región "bisagra" en donde el dinucleótido TA se repite en forma variable. En los alrededores de esta región la secuencia se conserva con distintos grados de homología.
- ♦ Una región 5' de alrededor 300 bases que constituye el cuerpo del elemento. Dentro de esta zona hay una región rica en G+C, sitio en donde la secuencia SIRE de SZ15 asociada al pseudogen ribosomal pierde completamente la homología.

ESTRUCTURA DEL ELEMENTO SIRE

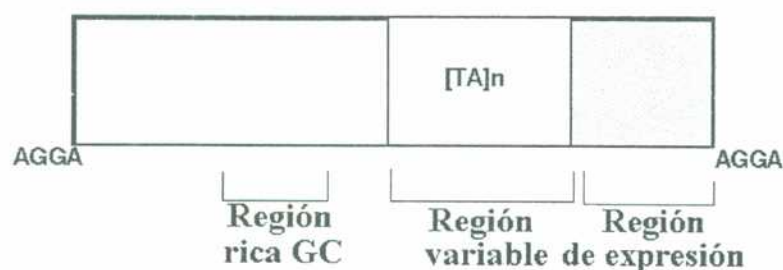
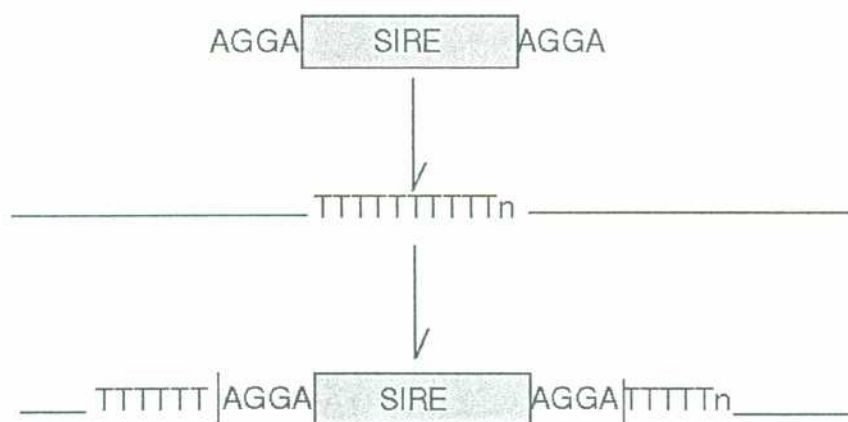


Figura 63 A.-
Estructura de un elemento SIRE típico.

Figura 63 B.-
Esquema de inserción
de un elemento SIRE en
el genoma de *T. cruzi*



5.3- Concepto de SAS (Sitios Asociados a SIRE): su importancia en el Proyecto Genoma de *Trypanosoma cruzi*.

Se ha demostrado que SIRE se encuentra en alto número de copias ubicado en diferentes puntos del genoma y representado en todos los cromosomas. Las regiones genómicas que se encuentran adyacentes al SIRE son por lo tanto secuencias no repetitivas, definen distintos loci en particular e incluso pueden ser secuencias que se encuentran representadas una sola vez en el genoma.

La posibilidad de generar eficientemente grupos de secuencias únicas o poco representadas que marquen segmentos específicos en el genoma es uno de los tópicos más importantes en el proceso del análisis de un genoma.

En la actualidad, se está desarrollando a nivel multinacional el proyecto de secuenciación del genoma de *Trypanosoma cruzi* y distintos laboratorios están llevando a cabo tareas definidas, una de esas tareas es la generación de STS (Sequence Tagged Sites). El STS por definición es una corta secuencia

de ADN que marca un región única en el genoma y esa región se encuentra perfectamente definida y caracterizada. Los STS ayudan en la elaboración de mapas genómicos de gran escala, paso previo y necesario para comenzar una secuenciación también en gran escala. Para el proyecto de genoma de *T. cruzi* se están desarrollando dos tipos de STS: derivados de microsatelites (86) y derivados del SIRE.

SIRE tiene un gran valor para el estudio del genoma, dado que rastreando SIRE se pueden aislar secuencias que marquen sitios únicos usando zonas ligadas a el. A estas secuencias le dimos en particular el nombre de SAS (SIRE Associates Sites o Sitios Asociados a SIRE).

Se purificaron distintos fragmentos asociados a SIRE utilizando los SZ. Estas regiones de ADN se caracterizaron por hibridización en Southern blots y PFGE. Adicionalmente estas experiencias completan la caracterización de SIRE mostrando la organización de los distintos loci que lo contienen.

Los datos se resumen en la tabla V:

TABLA V

| SZ | SAS derivado | SOUTHERN | PFGE |
|------|--------------|---------------|----------------|
| SZ2 | 2B-B360 | 5 | no determinado |
| SZ5 | 5B-B700 | 2 | 2 |
| SZ7 | 7B-B200 | 1 | 2 |
| SZ10 | 10P-P250 | al menos 5 | no determinado |
| SZ12 | 12E-E700 | gran cantidad | no determinado |
| SZ14 | 14B-B900 | 6 | 4 |
| SZ15 | 15E-P2000 | 5 | no determinado |
| SZ23 | 23P-P640 | 5 | 4 |
| SZ32 | 32B-B600 | 1 | no determinado |
| SZ36 | 36E-H1900 | 1 | no determinado |
| SZ39 | 39P-P1000 | 1 | 4 |

El SAS derivado de cada SZ se especifica con la siguiente convención: las letras indican las enzimas de restricción usadas para obtener el fragmento y los números que siguen indican el tamaño del mismo (B, BamHI; E, EcoRI;

P, PstI; H, HindIII). Delante de cada uno se ubica otro número que indica la procedencia SZ del fragmento.

Los números en las columnas de Southern y PFGE indican la cantidad de bandas detectadas usando cada SAS.

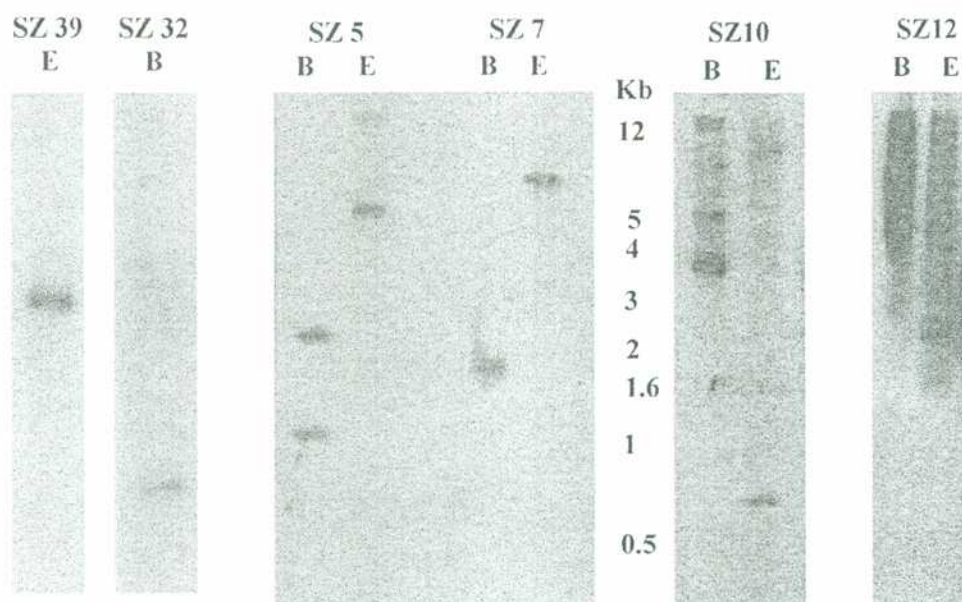
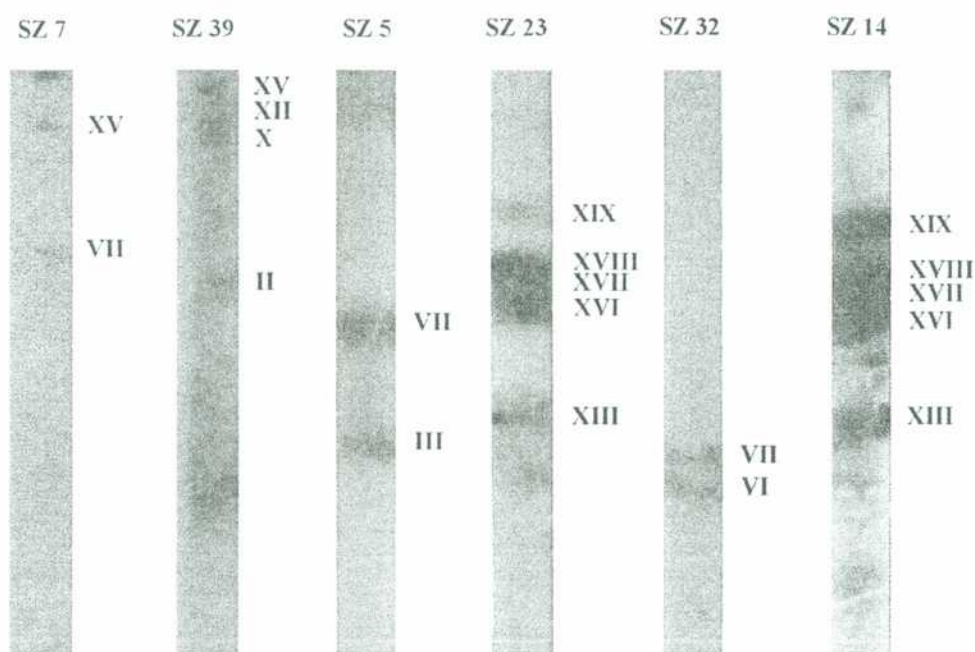
Los resultados de Southern blot (figura 64 A., página siguiente) muestran, por ejemplo, que el locus SZ10 está presente en más de una copia en el genoma (4 o 5 como mínimo) (obsérvese que la sonda no incluye la región de 130 pb común con otros clones descripta más arriba). SZ7, en cambio, muestra una sola banda de hibridación, indicando su presencia en una única ubicación genómica. Esto quiere decir que el entorno en donde SIRE se insertó es un locus único en el genoma. Al observar el PFGE (figura 64 B.) se detectan dos bandas, cabe la posibilidad de que se hayan detectado los cromosomas homólogos correspondientes. Para SZ5 se pueden observar dos bandas de hibridación en el Southern blot consistente con la presencia de dos loci genómicos para este entorno. Una hibridación adicional en PFGE muestra que también dos cromosomas distintos se marcan con la sonda SZ5, los cromosomas 7 (0.85 Mb) y 3 (0.7 Mb) (figura 64 B.), por lo tanto los dos loci SZ5 no están ligados en el genoma. La secuencia que se usó como sonda se puede ver en la figura 65 A.

La secuencia SAS derivada de SZ14 (figura 65 B y C.) mostró, al ser comparada en la base de datos del EBI, considerable homología con una familia particular de proteínas de superficie, GP85/Sialidasa/Neuraminidasa (87). A continuación se detalla esta homología:

TABLA VI

| NOMBRE | AUTOR | Nº ACC. EBI | HOMOLOGIA |
|---------------|----------------|-------------|-----------|
| TRBSIALB | KAHN,S. ET AL | M62736 | 54.3 % |
| TRBSA85KD | TAKLE,G. ET AL | J04667 | 55.4 % |
| TRB85KD | TAKLE,G. ET AL | M64836 | 55.4 % |
| TRBSIALA | KAHN,S. ET AL | M62735 | 53.4 % |
| TSA-1=85 | FOUTS,D. ET AL | S61039 | 55.5 % |
| TRBTSA1(PSEU) | TAKLE,G. ET AL | M91469 | 55.5 % |

Se muestra la homología con TRB85KD como ejemplo (figura 65 C.).

A**B****Figura 64-**

A. Southern blot de ADN genómico usando como sondas SAS purificados de los SZ indicados en la figura. Las características de cada SAS se puede ver también en la Tabla VI.. **B.** PFGE de cromosomas de *T. cruzi* hibridados con sondas SAS de los SZ indicados en la figura. Se indica el número del cromosoma detectado con cada sonda según Cano et al. (74)

A SZ5 (región 5' del inserto)-secuencia parcial

EcoRI

```

1  GAATTCAGGT  GGCCAACTGC  GTCCCTTGAT  GGTGCTACCC  ATCTACGCAT  CTCTCCACC
61  ACGAGAGCAA  AGGCGTATTT  ACGAGGTCCC  ACCACCAACC  ACACGGAAGG  TTGTGATCGC
121 GACAAACATT  GCTGAAACAT  CCATCACTAT  CGATGGAGTG  GTGTACGTCG  TGGATTGTGG
181 GCTTTGCAAG  CAAAACATTT  ATAATTACCA  GTCCATGGTG  GAGGAATTGC  GTGTGTTGCC
241 AATCTCGCAA  GCCAGCGCTA  AGCAGAGAAC  AGGTCGAGCA  GGACGAACTC  AAAGGGGAGT
301 GTTACCGTTC  GTACACGGTC  CACACCTTTC  GTAATGAGCT  TCCACAAGAA

```

B SZ14 (región 5' del inserto)-secuencia parcial

HindIII

```

1  GTCAACGCCC  AATTGAACCC  ACCATTTCGT  CCAAGCTTTC  GCTAACTTCC  CGCACGAAGA
61  AACGCCCTAT  TGCCGCTCTC  AATTGCCGTG  GGATTGAAAG  CGGACGCTAC  AACACCCTTT
121 ATATTGCTG  TGCTCAACAC  GATGTGTGTA  TTCCGGCCAT  TATTAACTC  ACCAAAAAAA
181 GACGACAAAT  TCCACGCAAA  AAGGATGGAA  CCCGCCAAAA  TATCCGCTGA  GAAAGCACAC
241 TCGGTGCCCC  ATCGCTATCA  TTAGCGTCGT  GTACACCCAT  CAAAGACAAC  CGTCAGTCAC
301 CAACACCCAC  AGAGCAGCAT  CCAACGTCTC  TCAGCC

```

SZ14 (zona SIRE)-secuencia parcial

```

1  TCCACACTGG  AAACCTTTTG  TTTGATGTCA  TCGGCTGATA  CTCCTACTCG  ATACTGGTGT
61  ACCTTTTCT  TCCAAAACCT  CGCAATTTCT  TTTTCAGCGT  CGGTTTCATT  ATTGGGATTG
121 GCGGGACCAC  ATCTCTGCCA  GCCAGTTAGC  GGCTTGTGTG  CTTTGTGCTG  CACATAAACA
181 ATCTCCACG  ACATGTCTCG  GGATAAGGCC  TCCCAATCAT  TAAATATGC  CGCCAGCTGT
241 TCAGTGAACA  CGTTCACAGT  ACTGGTTAGG  GTGTGGTGCT  CACTTAGCCG  TAGTGATCTG
301 CAGCCAACCA  TTATCTTCGG  GTT...

```

```

1...ACAGCACAAA  TTAGGAATTA  AACTCGTTT  TGTTTTTTTA  GTAAATTTA  GGAGCTTGTG
61  TGAAAAGAAT  GATCGCGGG  TAGCTGGCTA  ACTTAATTAA  TGTATGTGTT  TATCCTGATA
121 AATGAATGCA  TTCTTTATGG  TACTTTCTAC  CGTATGAATC  TTTTGGGAAG  AACGCGACTT
181 TGTAGGGGGC  GGGAGCCGAT  AGAGGCCGGA  TAATATTTTT  ATTTTATTTT  GCCATCCAC
241 CCACCCCTT  TTGATTCCCA  CCACGCGGCG  GGGTCT...

```

SIRE

C 3 TRB85KD <TRYPANOSOMA CRUZI 85 KD SURFACE ANTIGE 55.4 334 174 352 55.4% identity in 334 bp overlap, initial score: 174 optimized score: 352 COMPLEMENTARY key sequence

| | 10 | 20 | 30 | 40 | 50 |
|----------|--|----|----|----|----|
| SZ14.SEQ | GGCTGAGAGACGTTGGATGCTGCTCTGTGGGTGTTGGTGAC-TGACGGTTGTCTTTGATG | | | | |
| TRB85KD | GGGATTCATTGCTTCACCCCTCTTGTGCTAGTGCTGGTGGCGTAATTGCTGCCTTCGCCG | | | | |
| | 1060 1070 1080 1090 1100 1110 | | | | |
| SZ14.SEQ | GGTGT-ACACGA-CGCTAATGATAGCGATGGGGCACCAGAGTGCTTTCTCAGCGGATAT | | | | |
| TRB85KD | AAGGTCACCTGAGCTCCAAAATAAAGATAATAAATCAACTGAGC--CCTCTCTGATGC | | | | |
| | 1120 1130 1140 1150 1160 1170 | | | | |
| SZ14.SEQ | TTTGGCGGGTTCCATCCTTTTTCGCTGGAATTTGTCGTCCTTTTGGTGAGTTTAAT-- | | | | |
| TRB85KD | AGTTGCTTGGTACATTGACTCTGCGTGGGAATGGTCCACTCTTGTGGTGAGGTCAATAA | | | | |
| | 1180 1190 1200 1210 1220 1230 | | | | |
| SZ14.SEQ | -AATGGCCGGAATACACACATCGTGTGAGCACAGCGAATATAAAGG-GTGTGTAGCGT | | | | |
| TRB85KD | AAGTACATGGCAAGCACACCGTGTGGTAAAGTGGATGGAAAGGAGCGTTTGTATGT | | | | |
| | 1240 1250 1260 1270 1280 1290 | | | | |
| SZ14.SEQ | CCGCTTTCAATCCACGGCAATTGAGAGCGGCAATAGGGCGTTTCTTCGTGCGGGAAGTT | | | | |
| TRB85KD | -TGTGCTCCGCCCCACAACCACCAAGGACAATAAGGTGTTCTTTCGCGGGAAGCT | | | | |
| | 1300 1310 1320 1330 1340 1350 | | | | |

Figura 65-

A, Secuencia parcial de SZ5, usada como sonda en los ensayos de la figura 64.

B, Secuencia parcial de SZ14. Los puntos indican que faltan secuencias para interconectar los dos fragmentos. Se muestra la inserción de SIRE (recuadrado) en este locus.

C, Homología encontrada al rastrear el banco de datos EBI con SZ14. Se resume en tabla VI.

| | | | | |
|----------|--|------------------|------|-----------|
| | 300 | 310 | 320 | 330 |
| SZ14.SEQ | AGCGAAAGCTTGGTACGAATGGTGGGTTCA | ---ATTGGGCGTTGAC | | |
| | : | : | :::X | : |
| TRB85KD | CTGTTGCGTCCAATGTAAATGGGAGTTGGAGTCATGGCGGCTTGAA | | | |
| | 1360 | 1370 | 1380 | 1390 1400 |

La GP85 Sialidasa/Neuraminidasa de *T. cruzi* es una de las moléculas más importantes en la superficie del parásito ya que se supone tiene un rol preponderante en la penetración celular y evasión del sistema inmune (87).

Esta región de SZ14 se usó como sonda SAS en Southern blot y PFGE de cromosomas y los resultados concuerdan con esta homología, dado que la sonda identifica al menos 4 cromosomas distintos y la familia de GP85 es supernumeraria y distribuida por varios cromosomas (figura 64 A y B.).

Las restantes secuencias SAS purificadas no dieron homologías significativas con otras secuencias en las bases de datos.

En la actualidad, algunas de estas secuencias SAS ya se están utilizando en el rastreo de bancos genómicos en YACs y BACs.

5.4- Uso de SIRE en la diferenciación de cepas y en el proyecto genoma: Inter-SIRE PCR, SIRE-Bubble-PCR.

Muchas de las técnicas desarrolladas inicialmente para el proyecto genoma humano, se están transfiriendo a los proyectos de genomas de otros organismos.

Otra de las formas mas comunes de generación de STSs es el uso las secuencias repetitivas en reacciones de PCR, ALU-PCR en humanos o mas generalmente conocida como IRE-PCR (Interspersed Repetitive Element-PCR) (88).

Esta técnica hace uso de la propiedad repetitiva de estos elementos que pueden ubicarse mas de una vez por locus. Si dos elementos se encuentran relativamente cercanos y en direcciones opuestas, entonces el uso de un solo cebador que se una al extremo de la secuencia sería suficiente para lograr la amplificación del fragmento intergénico entre ellas (88).

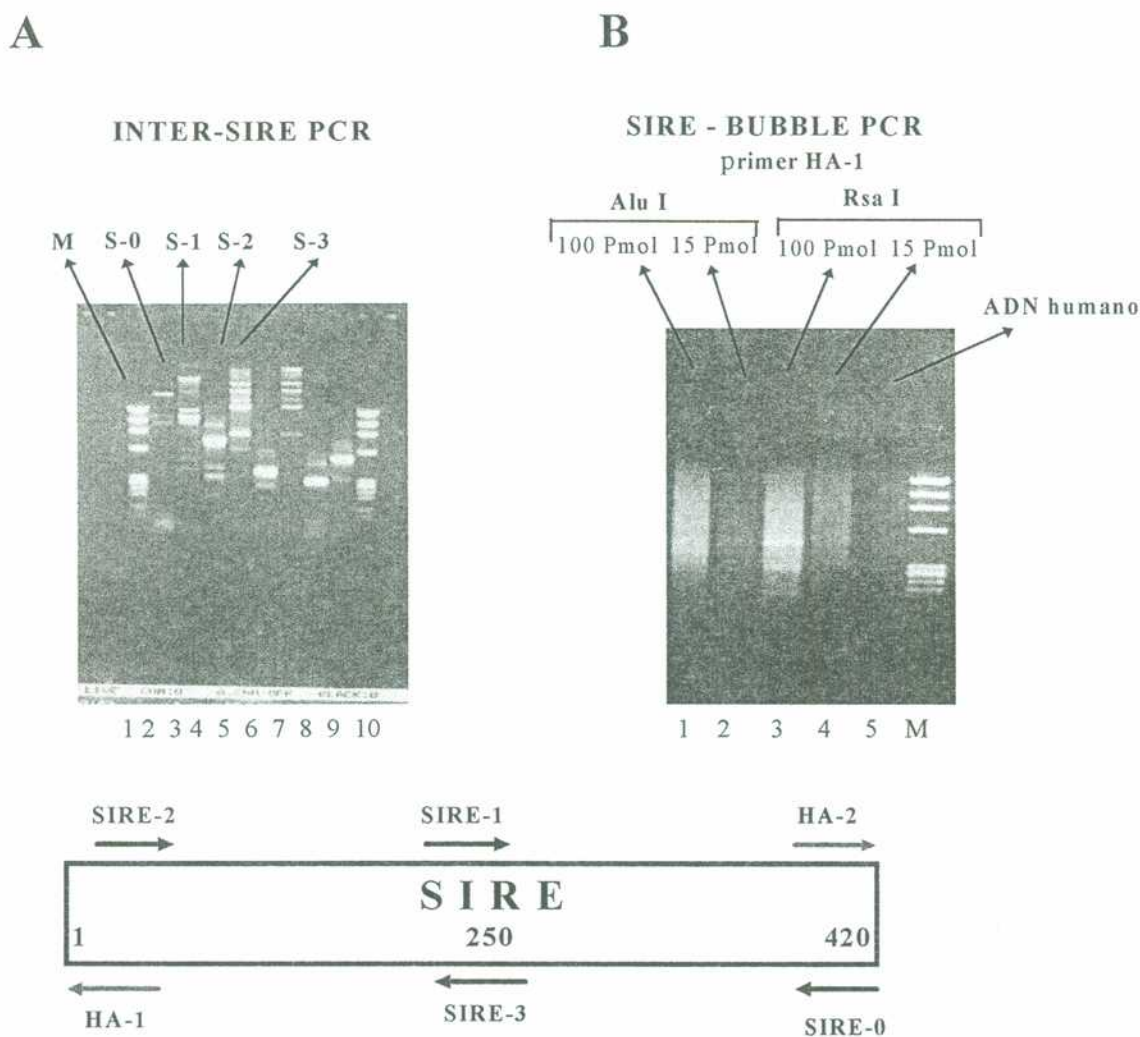
Una variación a la técnica de inter-IRE PCR es IRE-Bubble PCR (89). Inter-IRE depende enormemente de la distribución asimétrica del elemento dentro del genoma, y por lo tanto, deja poco representadas aquellas zonas con pobre contenido en secuencias IRE generando una distribución no uniforme de productos de PCR. IRE-Bubble se independiza de este efecto generando una distribución mas uniforme de productos. Básicamente el ADN genómico se corta con una enzima de restricción de corte frecuente y luego se ligan con extremos romos una serie de oligonucleótidos en los extremos de los fragmentos que servirán de anclaje para la amplificación de PCR usando otro cebador que se una al extremo de la secuencia IRE (el método es explicado en detalle en la sección materiales y métodos).

Ambos métodos los utilizamos usando la secuencia SIRE en *Trypanosoma cruzi* con dos objetivos:

- 1) Diferenciación de cepas (teniendo en cuenta que SIRE se distribuye de una forma particular en cada cepa)
- 2) Generación de SAS por PCR.

Los cebadores utilizados para las reacciones fueron en principio los mismos que ya se describieron en este capítulo (Sire-0, Sire-1, Sire-2, Sire-3). Las reacciones de PCR se llevaron a cabo sobre ADN genómico de la cepa CL colocando un solo cebador en cada reacción. Los resultados para estos cebadores se muestran en la figura 66 A. Se puede observar que todos producen un gran numero de bandas, particularmente Sire-1 y Sire-3, confirmando la distribución asimétrica de SIRE en el genoma. De esta forma, los fragmentos generados por PCR contienen SAS comprendidos entre dos elementos SIRE con dirección opuesta.

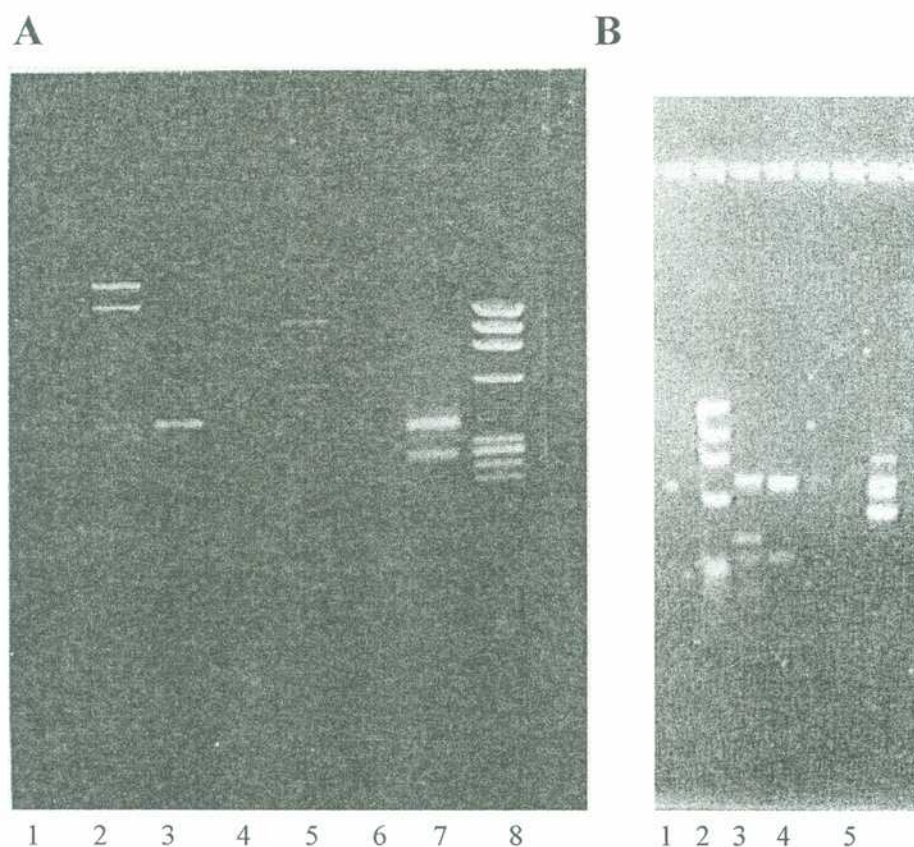
Cuando la reacción de PCR se combina con otro cebador específico de alguna de las secuencias SAS descritas anteriormente (Z7, cebador del SAS de SZ7 y Z2, cebador del SAS de SZ2), la reacción amplifica específicamente el SAS deseado (ver figura 66 A., calles 8 y 9).

**Figura 66-**

Comparación entre las reacciones de Inter-SIRE y SIRE-Bubble PCR. **A**, Inter-SIRE PCR. Templado, ADN genómico cepa CL Brenner. 1 y 10, marcadores ϕ X174 HaeIII. 2-5, Inter-SIRE con los cebadores especificados. 8 y 9, se incluyó además un cebador específico de locus (SZ2 y SZ7, respec.). **B**, SIRE-Bubble PCR.

Uno de los resultados más interesantes se obtiene de analizar el ADN genómico de varias cepas diferentes usando esta técnica con el cebador Sire-2. Las cepas CL, Tulahuen, Y, Maracay y G se diferencian claramente en los patrones generados por Inter-SIRE PCR (figura 67). De este modo, SIRE también puede ser importante en la caracterización de cepas.

Se diseñaron nuevos cebadores sobre los extremos de la secuencia SIRE, pero en este caso se dirigen hacia afuera del elemento (de hecho son

**Figura 67-**

Inter-SIRE en la identificación de cepas de T. cruzi. A, 1,2 y 3, cebador HA-1 con los templados Tulahuen, Maracay y cepa G, respec. 4, 5 y 6, cebador SIRE-1 (S-1) con los templados Tulahuen, Maracay y G, respec. 7, no se analiza. 8, marcador ϕ X174 HaeIII.. B, 1, Marcador ϕ X174 HaeIII. 2, Cepa CL. 3, Cepa Maracay. 4, Cepa Tulahuen. 5, Cepa G. El cebador utilizado fue SIRE-2.

antisentido de los cebadores Sire-0 y Sire-2) denominados HA-1 y HA-2 (ver figura 66). Se realizaron nuevas reacciones confirmatorias con estos cebadores, en donde se encontró que HA-1 diferenció mejor los patrones de las cepas que HA-2. HA-1 diferenció claramente las cepas Tulahuen, Maracay y G por este método (figura 67), confirmando que SIRE puede ser útil en la caracterización de cepas. Estas experiencias también muestran que el éxito en la caracterización de cepas depende del cebador elegido (observar las reacciones con HA-1, Sire-1 y Sire-2, en las figuras 66 y 67). De ahí surge la importancia de probar distintos cebadores en las reacciones para elegir el más adecuado, entendiendo por adecuado aquel cebador que logra diferenciar la mayor cantidad de cepas posibles.

Con el objetivo de lograr condiciones que permitieran obtener un mayor número de SAS en las reacciones de PCR, se utilizó la técnica de Bubble PCR explicada anteriormente.

Se probaron dos condiciones distintas: una cortando el ADN genómico con

AluI y otra cortando el ADN con RsaI. Para amplificar se utilizó el cebador de anclaje Bu-Primer (ver materiales y métodos para mas detalles) y el cebador de SIRE HA-1. Las reacciones de PCR para cada uno de los cortes se probaron en dos concentraciones distintas de templado que variaban en un orden (100 pmoles y 15 pmoles). Esto último es importante para optimizar la obtención de un mayor número de bandas.

Los resultados se pueden ver en la figura 66 B. Aun la condición con menor concentración de templado generó un complejo grupo de productos amplificados que aparece como un chorreado cuando el gel se tiñe con bromuro de etidio, sugiriendo que el número de productos amplificados es tan grande que obstaculiza la visualización de bandas individuales (observar que el control en el cual se omitió el agregado del cebador HA-1 no produjo amplificación alguna). Solo se pueden ver dos o tres bandas individuales intensas que posiblemente correspondan a zonas del genoma en las cuales SIRE se encuentra altamente representado.

Comparando con la reacción de inter-SIRE PCR (figura 66 A.), se puede ver que esta reacción es mucho mas sensible en la amplificación de secuencias SAS. Los productos generados por esta reacción se pueden clonar en un vector plasmídico para generar una biblioteca específica de SAS para la obtención de marcadores únicos de distintas regiones de cromosomas de *T.cruzi*.

También se puede aplicar directamente sobre las bibliotecas de YACs existentes de *T.cruzi* para generar todos los SAS posibles provenientes de un único cromosoma del parásito. Luego estos SAS pueden ser usados en el ordenamiento de los BACs o cósmidos necesarios para generar un mapa cromosómico. En estos casos todavía nos faltan realizar los controles necesarios sobre el ADN de levadura.

Las reacciones de Inter-SIRE PCR, de todas formas, son útiles para la caracterización de cepas, dado que al obtener menor número de bandas permite observar patrones distintivos para las distintas cepas.



A) El genoma de *T. cruzi* y la organización de genes: el ejemplo de TcP2 β y el resto de las proteínas P.

En eucariontes superiores la familia de proteínas P consta de un solo gen activo para cada tipo: P1, P2 y P0. La levadura *Saccharomyces cerevisiae* posee 5 genes de proteínas P: dos genes que codifican para dos tipos de la proteína P1, dos genes para dos tipos de P2 y un único gen P0 (8). El resto de los eucariontes inferiores en general siguen esta regla (9). El clonado de proteínas P en *T. cruzi* nos mostró que la situación en este parásito es mucho mas compleja. Existen al menos 4 genes que codifican para un tipo de proteína P2 (TcP2 β), dos genes que codifican para otro tipo de proteína P2 (TcP2 α), un único gen para P1 y dos genes P0 (90, 91). Esto suma al menos 9 genes que codifican para proteínas P en *T. cruzi* y todos se transcriben. Muchos de ellos presentan mínimas variaciones en la secuencia proteica, pero son diferenciables en cambios en el pI o variaciones de aminoácidos que podrían influir en la fosforilación de la proteína. De modo que todas podrían a ser funcionalmente distintas. Aun no hemos puesto a punto ensayos funcionales de estas proteínas como para determinar la importancia de estos cambios. Tampoco podemos determinar como interaccionan esta cantidad de proteínas P con los otros componentes de la estructura del ribosoma.

No se conocen con exactitud las diferencias entre el proceso de traducción en tripanosomas y el resto de los eucariontes, ni como influyen la inusual estructura de capping y el SL que llevan todos los ARNm del parásito. De todas formas, las proteínas P no se alejan de los datos que surgen a diario del clonado de genes en tripanosomas: proteínas y genes altamente conservados en toda la escala evolutiva suelen presentar profundos cambios en el orden de los kinetoplastidos

(proteínas P, Histonas (92), ARN ribosomales (93), etc.).

Caracterización general

Por otra parte, proponemos una nomenclatura unificada para denominar a las proteínas P de parásitos unicelulares que permite una comparación y manejo de la información mas adecuado.

El sistema P de *T. cruzi* se compone de las siguientes proteínas:

- TcP1, codificada por un gen.
- TcP2 α , codificada por al menos 2 genes ubicados en distintos loci genómicos.
- TcP2 β , codificada por al menos 5 genes ubicados en 4 loci genómicos diferentes.
- TcP0, codificada por 2 genes ubicados en distintos loci genómicos.

Este sistema P presenta las siguientes características diferenciales con respecto a los de otros organismos:

- Esta compuesto por al menos 9 genes activos en total para las distintas proteínas que lo componen. Los eucariontes inferiores (ej. Levadura) presentan 5 genes y los eucariontes superiores 3 (ej. humanos).
- Presenta 2 extremos C-terminales diferentes distinguibles por secuencia y por reactividad inmunológica (TcP1, TcP2 α y TcP2 β por un lado; TcP0 por el otro).
- Es el único organismo eucarionte que presenta un extremo C-terminal similar a de Archaeobacterias (TcP0).
- Es el único eucarionte inferior que contiene solo un tipo de proteína P1. El resto presenta 2 tipos (P1 α y P1 β).

TcP2 β se encuentra codificada por al menos 5 genes en 4 loci distintos y tiene niveles de expresión similares a los alcanzados por TcP1 codificada por un gen de copia única. Esto implica que el dosaje genico de TcP2 β es importante para alcanzar niveles altos de expresión. Los tripanosomas regulan la transcripción por mecanismos post-transcripcionales controlados principalmente por trans-splicing (53, 54, 55). Esto sugiere la importancia de comparar las secuencias 5' de los genes TcP2 β y TcP1 en busca de diferencias en las señales de trans-splicing que puedan explicar mejor como las regiones de polipirimidinas son capaces de ejercer regulación sobre los niveles finales de ARNm.

TcP2 α es codificada por al menos 2 genes y presenta los niveles mas bajos de expresión (5 veces menor que TcP1 por ejemplo). Estudios realizados en *Leishmania* (43) demuestran que los niveles de ARNm de un gen aumentan de acuerdo al grado de repetición de la región de polipirimidinas de la región 5'. Este trabajo se realizó usando construcciones artificiales y midiendo la expresión del gen CAT en ensayos de transfección. Una vez clonados los fragmentos genómicos conteniendo los otros genes P, el análisis comparativo de las regiones 5' de los mismos permitirá observar si esto es precisamente lo que ocurre en *T. cruzi*.

El análisis de la expresión en parásitos epimastigotes sometidos a golpes de calor (de 28°C a 37°C), simulando una de las condiciones del pasaje del insecto vector al hospedador mamífero, muestra muy pocas variaciones en los niveles de ARNm de las proteínas P. Solamente TcP1 presenta una pequeña disminución en los parásitos incubados a 37°C. Skeiky et al. (91) realizan mediciones similares pero directamente sobre tripomastigotes y epimastigotes de cultivo. Ellos observan una disminución más drástica en los niveles de ARNm de TcP1 en tripomastigotes, en tanto que los niveles del resto de las proteínas P se mantienen invariables (91). No es simple explicar esta posible regulación estadio específica únicamente de TcP1, teniendo en cuenta que las proteínas P se expresan constitutivamente o bien

varían su expresión coordinadamente solo durante periodos de "starvation" y "refeeding" en otros organismos estudiados (3).

B) El genoma de *T.cruzi* y las secuencias repetidas: el ejemplo de SIRE.

El genoma de *T.cruzi* es el más grande de los kinetoplastidos, 87 Mb. *T.brucei*, al igual que leishmania, tiene un genoma de aproximadamente 30 Mb. *T.cruzi*, a diferencia de *T.brucei* no posee minicromosomas, aunque presenta cromosomas pequeños como el I, 0.45 Mb y otros 7 por debajo de 1 Mb (74).

La abundancia de las secuencias repetitivas en los genomas eucarióticos ha estimulado muchas discusiones acerca de la función que pueden cumplir. Como se han encontrado por delante o por detrás de genes que codifican para proteínas se sugirió que podrían tener algún rol en la modulación de la expresión de los genes. La existencia de estas secuencias repetidas también se ha interpretado en el contexto de la potencialidad de los eucariotas para desarrollar estrategias que les permitan reaccionar ante cambios comparativamente rápidos, dándole plasticidad al genoma. Por otra parte, se ha sugerido también que estos elementos no tiene otra función mas que sobrevivir dentro de los genomas (94, 95).

En *T.cruzi* se han informado varias secuencias repetitivas que comprenden una porción significativa de su genoma, pero al igual que en otros organismos su función es desconocida.

SIRE y trans-splicing

SIRE comparte algunas características con estas secuencias repetidas: la dispersión en el genoma y la no interrupción de secuencias codificantes. Sin embargo, SIRE parece ser único en el hecho de que su inserción delante del gen TcP2 β -H1.8 introduce un sitio aceptor SL funcional conjuntamente con una región polipirimidinica. Esto resulta de gran importancia de acuerdo al

modelo propuesto de “scanning” para la selección del sitio de poliadenilación (44). En efecto, el sitio de reconocimiento del SL es el que determina la región de poliadenilación del ARNm anterior en un ARN policistrónico (ver introducción). LeBowitz et al. demuestran que en *Leishmania major* la poliadenilación del gen inmediatamente anterior es dictada por la región de trans-splicing del gen posterior, sin embargo esta poliadenilación no ocurre en cualquier sitio, sino dentro de una “ventana” permitida rica en regiones (T)nA. Esto hace que no exista una señal consenso de poliadenilación y que el sitio de poliadenilación no sea tan definido como en los eucariontes superiores en donde la señal AAUAAA define exactamente el sitio. En tripanosomas cualquier sitio dentro de esa “ventana” puede ser elegido para agregar la cola de poliA. Esto concuerda con nuestros resultados de clonado de ARNm de TcP2 β . Todos los ARNm TcP2 β se poliadenilaron en sitios diferentes pero cercanos. Esta dispersión del sitio de poliadenilación (ver figura 33, del capítulo 4 de resultados) siempre ocurre dentro de una ventana rica en secuencias (T)nA, de acuerdo con lo demostrado por LeBowitz et al para *Leishmania major*, indicando entonces que la región de trans-splicing del gen posterior esta marcando la ventana de poliadenilación.

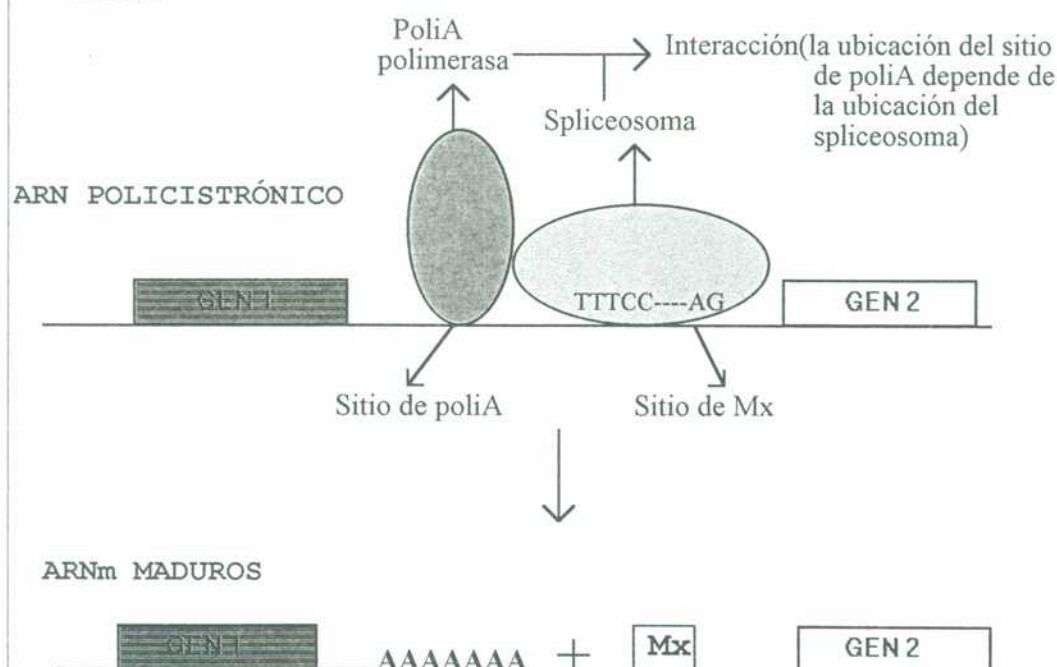


Figura 68-
Diagrama que muestra el acoplamiento entre la poliadenilación y el trans-splicing en tripanosomas.

Matthews et al. (96) demostraron que en *Trypanosoma brucei* una región común rica en pirimidinas es la que dirige las reacciones de trans-splicing y poliadenilación en la región intergénica de los genes de tubulina. Este hallazgo y los de LeBowitz et al permiten generalizar que las reacciones de trans-splicing y poliadenilación se encuentran acopladas en kinetoplastidos y que dependen de un elemento funcional común.

Es probable que un factor compartido funcione uniendo la región de pirimidinas para dirigir las reacciones. Matthews et al. especulan que el SL, que cumple a la vez funciones de U1 en tripanosomas, podría estar cumpliendo esta función dual en la poliadenilación y el trans-splicing (96).

Cuando SIRE (428pb) se inserta en una región intergénica, alarga la distancia entre dos genes, destruye la señal de trans-splicing original y presenta una nueva al gen que se encuentra hacia abajo

Por lo tanto, sugerimos que SIRE puede donar simultáneamente un nuevo sitio de poliadenilación a la región codificante ubicada hacia arriba de la inserción, de acuerdo a lo explicado anteriormente.

Más aún, el SIRE podría funcionar como “espaciador” del policistrón, donando un sitio aceptor de SL al final del mismo y dirigiendo la poliadenilación de la última región codificante. Lebowitz et al. ya plantean en su trabajo (44) la hipotética existencia de estos “espaciadores” al final de una unidad transcripcional para mantener la compatibilidad de su modelo. Matthews et al. especulan que alguna característica especial de la secuencia se encarga de marcar al gen terminal de un policistrón para su procesamiento por poliadenilación. La presencia de SIRE al final de una unidad transcripcional podría inducir la poliadenilación río arriba donando estas señales.

La visión actual sobre la modulación de la expresión en estos organismos es que los niveles finales de ARNm se obtienen a través de mecanismos de regulación post-transcripcional que involucran loops regulatorios en el pre-ARNm en combinación con tasas diferenciales de trans-splicing y

poliadenilacion que afectan su vida media (51, 52, 53). De allí la importancia de la presencia de la región pirimidinica funcional y del loop que se forma en la región 3' transcrita dentro del SIRE.

SIRE y su relación con otras secuencias repetidas de *T.cruzi*

Una secuencia repetida descrita por Requena et al. (97), E12, muestra en su parte 3' homologia con SIRE. En efecto, las ultimas 120 bases de E12 tienen una alta homología con las primeras 120 de otro elemento repetido, E13 (81) y con las ultimas 120 bases de SIRE. El resto de la secuencia de los elementos no tienen relación entre si ni de homologia ni de organización. El elemento E12, al igual que SIRE participa de un evento de transcripción. La región 5' de E12 se transcribe como parte de la región 3' no codificante de un ARNm cuyo producto aun no se ha identificado, y la poliadenilación de ese mensajero ocurre dentro de la secuencia transcrita de E12 (98). De modo que la participación de las secuencias repetidas en la modulación de la expresión de genes en tripanosoma, a nivel transcripcional o post-transcripcional, podría generalizarse

Una pregunta que queda sin resolver por el momento es el tipo de relacion existente entre elementos como E12, E13 y SIRE, que comparten una corta zona común de 120 bases en distintas posiciones en los 3 elementos (figura 43, capítulo 4 de resultados). Aun no podemos determinar si todas ellas tuvieron un origen común o adquirieron esta secuencia por recombinacion desigual. Sin embargo, lo que resulta interesante es que estos elementos puedan tener alguna relación entre si, dado que entre todos suman un porcentaje importante del genoma de *T.cruzi*. Teniendo en cuenta que todos los genes descritos hasta ahora se agrupan muy cerca unos de otros, separados por cortas regiones intergénicas, es probable que el resto del genoma se encuentre poblado de estas secuencias repetidas. Esto, quizás, se corresponda con el tamaño mayor del genoma de *T.cruzi* en relación a otros

kinetoplastidos.

Tanto E13, E12, E22 (97) como SIRE se encuentran repetidas en un rango de 10^3 - 10^4 copias dispersas por el genoma. Sin embargo SIRE es el elemento que parece presentar mayor dispersión de todos los mencionados. Observar la figura 69, en donde se comparan los southern blots de SIRE, E12 y E22, en particular la calle 2 digerida con BamHI, muestra que SIRE presenta una distribución mas homogénea en el genoma que los otros elementos. Este dato es importante al utilizar a SIRE como base para generar marcadores STSs en el proyecto genoma de *Trypanosoma cruzi*.

Figura 69. -
Southern blots compa-
rativos entre secuencias
repetidas de *T.cruzi*.
1, EcoRI. 2, BamHI. 3,
HindIII. 4, PstI. 5, Sall



Cuál es la causa por la cual un genoma relativamente pequeño como el de este organismo conserva tantas secuencias repetidas aun es un misterio. Quizás le brindan al parásito una plasticidad a su genoma que ayude a adaptarse a los cambios violentos a los cuales esta sometido. Parte de su habilidad para sobrevivir a cambios rápidos puede basarse en la adaptabilidad que le otorga la plasticidad de su genoma.

SIRE y la duplicación de genes

Otra función atribuible a SIRE es una función en la duplicación de genes. Observar la particular organización del locus H1.8: SIRE-TcP2 β -SIRE-TcP2 β -SIRE, y tener en cuenta la presencia dentro del elemento repetitivo de una putativa secuencia recombinógena Chi. Esta secuencia Chi, que ha sido descrita en diversas regiones del genoma de tripanosoma (*T.brucei* y *T.cruzi*), estaría involucrada en eventos de recombinación y se la encuentra en las proximidades de genes duplicados (99, 100). Chi ya fue asociada con otra secuencia repetida de *T.cruzi* que no es SIRE. En efecto, Novak et al (79), describen una secuencia medianamente repetida que se asocia al "cluster" de genes de ARN ribosomal, la familia SRE-1-2-3, en donde también encuentran que una secuencia Chi (GCTGGTGG) se asocia al elemento repetido. Secuencias tipo Chi también se encontraron en las cercanías de los genes de aldolasa, fosfoglicerato kinasa y gliceraldehido-3-fosfato deshidrogenasa, y en una secuencia tipo retrotransposon (99, 100). Los genes de estas enzimas glicolíticas con secuencias Chi en sus vecindades se encuentran organizados en tandems directos, con secuencias idénticas o prácticamente idénticas de sus genes. Se ha propuesto que estas secuencias tipo Chi podrían estar funcionando como putativos mediadores en el proceso de conversión génica en tripanosomas. Mas aun, recientemente se ha demostrado que elementos circulares extracromosomicos de *Leishmania mexicana* pueden ser formados por un "cross-over" entre secuencias tipo Chi (101).

Estos datos no difieren demasiado de lo observado por nosotros donde SIRE se asocia con los genes TcP2 β . SIRE lleva una secuencia Chi en uno de sus extremos y los genes TcP2 β se encuentran duplicados en tandem directo en torno a estas secuencias SIRE. Las secuencias de los genes TcP2 β en tandem son idénticas.

La asociación de Chi con secuencias repetidas como SIRE o SRE-1-2-3 es

aun mas importante, dado que estos elementos se encuentran dispersos por el genoma y , en el caso de SIRE, presente en todos los cromosomas.

Estructura y organización de SIRE

El trabajo presentado en esta tesis sobre SIRE presenta ciertas ventajas con respecto a otros elementos descritos. Su observación en varios loci en donde se encuentra ubicado nos permitió definir exactamente cual es su blanco de inserción y la estructura primaria del elemento (figura 63, capítulo 5 de resultados). SIRE se encuentra siempre ubicado en una región de polidT y teniendo en cuenta que las regiones ricas en T se encuentran ampliamente distribuidas en la zonas intergénicas de *T.cruzi*, parece poco probable que SIRE podruzca inactivaciones insercionales de genes.

Con respecto a la estructura del mismo, podemos definir tres regiones:

- La región con señales de trans-splicing que se transcribe (I).
- una región de (TA)_n variable. Alrededor de esta región la homología es diversa (II).
- una región 5' mas conservada con una zona rica en GC (III).

Es notable la conservación de la región I aun cuando la zona adyacente presente homología variable entre los distintos SIRE. Esto esta marcando la importancia de esa región que contiene las señales de trans-splicing y la presión de selección por mantenerla.

Aun no sabemos si SIRE se encuentra también formando parte de un elemento mayor del tipo de retrotransposones. Es conocido que las regiones LTR de los retrotransposones tienen la capacidad de moverse en forma autónoma y promover la transcripción de genes que se ubican en sus proximidades. Podría ser que SIRE formara parte o fuera un LTR en si mismo, pero no tenemos más elementos para afirmarlo.

En conclusión, el estudio de los elementos repetidos en *T.cruzi* está en la actualidad aportando datos valiosísimos sobre la estructura y plasticidad del genoma de este parásito.

La siguiente tabla resume las secuencias repetitivas mas importantes descriptas hasta la fecha en *T.cruzi*:

| NOMBRE | N DE COPIAS | CROMOSOMAS | CITA |
|--------------|----------------|------------|------------------------|
| Minisatelite | 10^5 | Al menos 8 | 102 |
| E13 | $5 \cdot 10^4$ | La mayoría | 81 |
| E12 | $5 \cdot 10^3$ | La mayoría | 97 |
| E22 | $5 \cdot 10^3$ | La mayoría | 97 |
| SRE1-2-3 | $5 \cdot 10^2$ | Pocos | 79 |
| DFG | $2 \cdot 10^2$ | Pocos | 103 |
| L1TC | $2 \cdot 10^3$ | -- | 104 |
| SIRE | $2 \cdot 10^3$ | Todos | 69 y la presente tesis |

SIRE y el Proyecto Genoma de *Trypanosoma cruzi*.

La capacidad de generar rápida y eficientemente un grupo de secuencias cortas de copia única dirigidas a un segmento específico del genoma es uno de los tópicos mas importantes en el proceso de análisis genómico. El desarrollo de la técnica de PCR basada en elementos repetitivos dispersos (IRE-PCR) a tenido un gran impacto en el avance del proyecto Genoma Humano. IRE-PCR hizo posible la amplificación específicas de secuencias de ADN humanas contenidas en células somáticas híbridas, cromosomas artificiales de levadura (YAC), cosmidos, etc. IRE-PCR es particularmente poderosa porque puede ser aplicada sin conocimiento previo de la secuencia contenida dentro del segmento genómico que se está analizando. Sin embargo IRE-PCR presenta limitaciones. IRE-PCR funcionara únicamente entre pares de secuencias IRE que estén a una distancia y orientación apropiada una respecto a la otra; incluso se ve restringida por la aparente distribución asimétrica de la secuencia repetida dentro del genoma. Este particionamiento dela repetición, en donde puede haber regiones del genoma ricas o pobres en contenido de la secuencia IRE, llevan a una distribución no uniforme de los productos IRE-PCR dejando algunas regiones del genoma significativamente poco representadas (88, 89).

El desarrollo de la técnica de IRE-Bubble PCR supera estas limitaciones y extiende la utilidad básica de esta estrategia en el análisis de genomas. Con IRE-Bubble PCR virtualmente cualquier fragmento de ADN que contenga un único IRE puede ser amplificado por esta técnica, dado que se basa en el anclaje de un “linker” independiente de la secuencia analizar.

El análisis que se ensayó en el marco de la presente tesis se basó en usar SIRE como elemento repetido para aplicar estas técnicas. SIRE ha probado ser especie específico y permitió diferenciar distintas cepas en ensayos de PCR, usando la técnica de inter-SIRE PCR. Por otra parte SIRE presenta ciertas ventajas sobre otras secuencias repetidas de *T. cruzi* ya que tiene una distribución más uniforme en el genoma y presencia en todos los cromosomas (ver figura 69), cualidades de gran importancia para utilizar las técnicas de inter-SIRE y SIRE-Bubble PCR para obtener STSs (SAS) para la construcción del mapa físico del genoma nuclear de *Trypanosoma cruzi*.

En la actualidad, se ha realizado el Kariotipo molecular de la cepa CL Brener elegida para el proyecto genoma y se determinó que contiene 20 cromosomas entre 0.45 Mb y 3.50 Mb, y un genoma nuclear de 87 Mb (74). Nuestro laboratorio ha decidido clonar y secuenciar el cromosoma I (0.45 Mb) en su totalidad. SIRE presenta una intensa hibridación con este cromosoma (figura 48). De modo que la técnica de SIRE-Bubble puede ser aplicada sobre los distintos BACs que cubran regiones del cromosoma I para obtener SAS que permitan el aislamiento de nuevos BACs hasta completar el mapa físico.



- Se obtuvieron y caracterizaron las secuencias completas de los ARNm de TcP1 y TcP2 α .
- Se obtuvieron y caracterizaron nuevas variantes de ARNm de TcP2 β .
- Se demostró que TcP1 es codificado por un solo gen, TcP2 α al menos por 2 genes en loci diferentes y TcP2 β al menos por 5 genes distribuidos en 4 loci genómicos diferentes (H6.4, H1.8, H1.5, H1.3).
- La transcripción y traducción in vitro de las proteínas caracterizadas demuestra que existe una proteína P de 17 kd que aun no ha sido clonada.
- La abundancia relativa de los ARNm de proteínas P es la siguiente: TcP1 0.1%, TcP2 β 0.15% y TcP2 α 0.02%.
- Por lo expuesto en 3 y 5, las señales reguladoras de transcripción de TcP1 deben ser las más fuertes de los 3 genes estudiados.
- Una secuencia repetida (SIRE) se asocia al tandem de genes TcP2 β en el locus H1.8.
- SIRE contiene una señal funcional de trans-splicing y afecta la expresión de TcP2 β en el locus H1.8.
- SIRE está presente en todos los cromosomas de *T. cruzi*.
- SIRE se inserta siempre en una región rica en T.
- La distribución de SIRE es especie específica (ejemplo, cepa G).
- Los sitios asociados a SIRE (SAS) son de gran importancia en la generación de STSs para el proyecto genoma de *Trypanosoma cruzi*.
- los SAS también son importantes en la diferenciación y caracterización de cepas.

1. Kirshom, L. and Isaksson, L. (1985) Involvement of ribosomal protein L7/L12 in control of transcriptional accuracy. **Proc. Natl. Acad. Sci. USA** 82, 717-721.
2. Walieczech, J. et al. (1989). Comparative cross-linking study on the 50S ribosomal subunit from *Escherichia coli*. **Biochemistry** 28, 4099-105.
3. Zinker, S. and J. R. Warner. (1976) The ribosomal proteins of *Saccharomyces cerevisiae*. **J. Biol. Chem.** 251, 1799-1807.
4. Uchiumi, T. et al. (1987). Topography and stoichiometry of acidic proteins in large ribosomal subunits from *Artemia salina* as determined by crosslinking. **Proc. Natl. Acad. Sci. USA** 84, 5580-5584.
5. Saenz Robles, M. et al. (1990). **Biochimica et Biophysica Acta**. 1050, 51-55.
6. Amons, R., Pluijms, W. y Holler, W. (1979) The primary structure of ribosomal protein eL12/eL12-P from *Artemia salina* 80S ribosomes. **FESS Lett.** 104, 85-89.
7. Rich, B.E. and Steitz, J. (1987) Human acidic ribosomal proteins P0, P1 and P2: analysis of cDNA clones, in vitro synthesis and assembly. **Mol. Cell. Biol.** 7, 4065-4074.
8. Remacha, M., Saenz-Robles, M.T., Dolores Vilella, M. and Ballesta, J.P.G. (1988) Independent genes coding for three acidic proteins of the large ribosomal subunit from *Saccharomyces cerevisiae*. **J. Biol. Chem.** 263, 9094-9101.
9. Beltrame, M. and M. Bianchi. (1990) A gene family for acidic ribosomal proteins in *Schizosaccharomyces pombe*: two essential and two nonessential genes. **Mol. Cell. Biol.** 10, 2341-2348.
10. Prieto, J, Candel, E. and Coloma, A. (1991) Nucleotide sequence of a cDNA encoding ribosomal acidic phosphoprotein P1 from *dictiostelium discoideum*. **Nucleic Acids Res.** 19, 1340.
11. Prieto, J, Candel, E. and Coloma, A. (1991) Nucleotide sequence of a cDNA encoding ribosomal acidic phosphoprotein P2 from *dictiostelium discoideum*. **Nucleic Acids Res.** 19, 1341.

12. Prieto, J, Candel, E. and Coloma, A. (1991) Nucleotide sequence of a cDNA encoding ribosomal acidic phosphoprotein P1 from *dictyostelium discoideum*. **Nucleic Acids Res.** 19, 1342.
13. Wigboldus, J. (1987) cDNA and deduced aminoacid sequence of Drosophila rp21C, another "A" type ribosomal protein. **Nucleic Acids Res.** 15, 10064.
14. Ramirez, C.; Shimmin, L.; Newton, H; Matheson, T.; and Dennis, P. (1989) Structure and evolution of the L11, L1, L10 and L12 equivalent ribosomal proteins in eubacteria, archaebacteria and eucariontes. **Can. J. Microbiol.** 35, 234-244.
15. Liljas, A., Thirup, S. and Matheson, A.T. (1985) Evolutionary aspects of ribosome-factor interactions. **Chemica Scripta** 26B, 109-119.
16. Hunter Newton, C., Shimmin, L.C., Yee, J. and Dennis, P.P. (1990) A family of genes encode the multiple forms of the *Saccharomyces cerevisiae* ribosomal proteins equivalent to the *Escherichia coli* L12 protein and a single form of the L10-equivalent ribosomal protein. **J. Bacteriol.** 172, 579-588.
- 17- Shimmin, L. and Newton, H. (1989) Organization of genes encoding the L11, L1, L10 and L12 equivalent ribosomal proteins in eubacteria, archaebacteria and eucaryotes. **Can. J. Microbiol.** 35, 164-170.
18. Naranda and J.P.G. Ballesta. (1991) Phosphorilation controls binding of acidic proteins to the ribosome. **Proc. Natl. Acad. Sci.**, 88, 10563-10567.
19. Santos, C.; Reyes-Ortiz, B.; Naranda, T.; Remacha, M.; Ballesta, J.P.G. (1993) The acidic phosphoproteins from *Saccharomyces cerevisiae* ribosomes. NH2-Terminal acetylation is a conserved difference between P1 and P2 proteins. **Biochemistry**, 16, 4231-4236.
20. T. S. Hansen, P. H. Andreasen, H. Dreisig, P. Hojrup, H. Nielsen, J. Engberg & K. Kristiansen (1991) Tetrahymena thermophila acidic ribosomal protein L37 contains an archaebacterial type of C-terminus. **Gene** .105, 143-150.
21. Hoare, C. (1972) The trypanosomes of mammals. Blackwell Scientific Publications
22. Miles, M. and Cibulskis, R. (1986) Zymodeme characterization of *Trypanosoma cruzi*. **Parasitol. Today**. 4, 94-101.

23. Cura, E.; de Titto, E.; Segura, E. (1992) Actualizaciones en la enfermedad de Chagas. Simposio satelite.
24. Kierzembaum, F. (1986) Autoimmunity in Chagas' disease. **J. Parasitol.** 72, 201-211
25. Ferrari, I.; Levin, M.; Wallukat, G.; Elise, R.; Lebesgue, D.; Chiale, P.; Elizardi, M.; Rosembaum, M.; Hoebeke, J. (1995) Molecular mimicry between the immunodominant ribosomal protein P0 of *Trypanosoma cruzi* and a functional epitope on human β 1-Adrenergic receptor. **J. Exp. Med.** 182, 59-65.
26. Elkon, K.; Bonfa, E.; Llobet, R.; Duho, W.; Weissbach, H.; Brot, N. (1988) Properties of the ribosomal P2 protein autoantigen are similar to those of foreign proteins antigens. **Proc. Natl. Acad. Sci.** 85, 5186-5189.
27. Elkon, K.; Parnassa, A.; Lee Foster, C. (1985) Lupus autoantibodies target ribosomal P proteins. **J. Exp. Med.** 162, 459-471.
28. Kerner, N.; Liégeard, P.; Levin, M.; Hontebeyrie-Joskowicz, M. (1991) *Trypanosoma cruzi*: antibodies to a MAP-like protein in Chronic Chagas' disease cross-react with mammalian cytoskeleton. **Exp. Parasitology.** 73.
29. Levin, M.J., Mesri, E., Benarous, R., Levitus, G., Schijman, A., Levy-Yeyati, P., Chiale, P.A., Ruiz, A.M., Kahn, A., Rosenbaum, M.B., Torres, H.N. and Segura, E.L. (1989) Identification of major *Trypanosoma cruzi* antigenic determinants in chronic Chagas heart disease. **Am. J. Trop. Med. Hyg.** 41, 530-538.
30. Mesri, E.A., Levitus, G., Hontebeyrie-Joskowicz, M., Dighiero, G., Van Regenmortel, M.H.V. and Levin, M. (1990) Major *Trypanosoma cruzi* determinant in Chagas heart disease shares homology with the Systemic Lupus Erythematosus ribosomal P protein epitope. **J. Clin. Microbiol.** 28, 1219-1224.
31. Levitus, G., Hontebeyrie-Joskowicz, M., Van Regenmortel, M.H.V. and Levin, M.J. (1991) Humoral autoimmune response to ribosomal P proteins in chronic Chagas heart disease. **Clin. Exp. Immunol.** 85, 413-417.
32. Cross, G. (1990) Cellular and genetic aspects of antigenic variation in Trypanosomes. **Ann. Rev. Immunol.** 8, 83-110.

33. Kendall, G.; Wilderspin, A.; Ashall, F.; Miles, M.; Kelly, J. (1990) *T.cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase does not conform to the hotspot topogenic signal model. **EMBO J.** 9, 2751-2758.
34. Marchand, M.; Poliszczak, A.; Gibson, W.; Wierenga, R.; Oppendoes, F.; Michels, P. (1988) Characterization of the genes for fructose-biphosphate aldolase in *T.brucei*. **Mol. Biochem. Parasitol.** 29, 65-76.
35. Requena, J.; Lopez, M.; Jimenez-Ruiz, A.; de la Torre, J.; Alonso, C. (1988) A head to tail tandem organization of HSP70 genes in *Trypanosoma cruzi*. **Nucleic Acids Res.** 16, 1393-1406.
36. Roditi, I.; Pearson, T. (1990) The procyclin coat of african trypanosomes. **Parasitol. Today** 6, 79-82.
37. Muhich, M. and Boothroyd, J.(1988) Polycistronic transcripts in Trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. **Mol. Cell. Biol.** 8, 3837-3846.
38. Agabian, N. (1990) Trans-splicing of nuclear pre-mRNAs. **CELL.** 61, 1157-1160.
39. Laird, P.; Zomerdijk, J.; de Korte, D. and Borst, P. (1987) In vivo labelling of intermediates in the discontinuous synthesis of mRNA in *Trypanosoma brucei*. **EMBO J.** 4, 1055-1062.
40. Walder, J.; Eder, P.; Engman, D.; Brentano, R.; Walder, R.; Knutzon, D.; Dorfman, D.; Donelson, J. (1986) The 35-nucleotide spliced leader sequence is common to all trypanosome messenger RNAs. **SCIENCE.** 233, 569-571.
41. Bruzik, J.; Steitz, J. (1990) Spliced leader RNA sequences can substitute for the essential 5' end of U1 RNA during splicing in a mammalian in vitro system. **CELL** 62, 889-899.
42. McNally, K.; Agabian, N. (1992) *Trypanosoma brucei* spliced-leader RNA methylations are required for trans splicing in vivo. **Mol. Cell. Biol.** 12, 4844-4851.
43. Curotto de Lafaille, M.; Laban, A.; Wirth, D. (1992) Gene expression in Leishmania: Analysis of essential 5' DNA sequences. **Proc. Natl. Acad. Sci.** 89, 2703-2707.

44. LeBowitz, J.H., Smith, H.Q., Rusche, L. and Beverley, M. (1993) Coupling of poly(A) site selection and trans-splicing in *Leishmania*. **Genes & Development** 7, 996-1007.
45. Conrad, R.; Liou, R. F.; Blumenthal, T. (1993) Conversion of a trans-splicing gene into a conventional gene by introduction of a splice donor site. **EMBO J.** 12, 1249-1255.
46. Van Der Ploeg, L.H.T.; Gottesdiener, K.; Lee, M.G.S (1992) Antigenic variation in african trypanosomes. **Trends Genet.** 8, 452-457.
47. Gottesdiener, K.; Chung, H.; Brown, S.; Lee, M.G.S.; Van Der Ploeg, L.H.T. (1991) Characterization of VSG gene expression site promoters and promoter-associated DNA rearrangement events. **Mol. Cell. Biol.** 11, 2467-2480.
48. Johnson, P.; Kooter, J.; Borst, P. (1987) Inactivation of transcription by U.V. irradiation of *T.brucei* provides evidence for a multicistronic transcription unit including a VSG gene. **CELL** 51, 273-281.
49. Evers, R.; Hammer, A.; Kock, J.; Jess, W.; Borst, P.; Memet, S.; Cornelissen, A. (1989) *Trypanosoma brucei* contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. **CELL** 56, 585-597.
50. Rudenko, G.; Lee, M.G.S.; Van Der Ploeg, L.H.T. (1992) The PARP and VSG genes of *trypanosoma brucei* do not resemble RNA polymerase II transcription units in sensitivity to sarkosyl in nuclear run-on assays. **Nucl. Acids Res.** 20, 303-306.
51. Janz, L. and Clayton, C. (1994) The PARP and rRNA promoters of *Trypanosoma brucei* are composed of dissimilar sequences that are functionally interchangeable. **Mol. Cell. Biol.** 14, 5804-5811.
52. Chung, H.; Lee, M.G.S.; Dietrich, P.; Huang, J.; Van Der Ploeg, L.H.T. (1993) Disruption of largest subunit RNA polymerase II genes in *Trypanosoma brucei*. **Mol. Cell. Biol.** 13, 3735-3743.
53. Pays, E.; Coquelet, H.; Tebabi, P.; Pays, A.; Jefferies, D.; Steinert, M.; Koenig, E.; Williams, R. Roditi, I. (1990) *Trypanosoma brucei*: constitutive activity of the VSG and procyclin gene promoters. **EMBO J.** 9, 3145-3151.

54. Hehl, A.; Roditi, I. (1994) The regulation of procyclin expression in *Trypanosoma brucei*: making or braking the rules?. **Parasitol. Today** 10, 442-445.
55. Graham, S.V. (1995) Mechanisms of stage regulated gene expression in kinetoplastida. **Parasitol Today** 11, 217-223.
56. Cech, T. (1991) RNA Editing: world's smallest introns? **CELL** 64, 667-669.
57. Oppendoes, F. (1987) Compartmentation of carbohydrate metabolism in Trypanosomes. **Ann. Rev. Microbiol.** 41, 127-151
58. Perez-Morga, D.; Englund, P. (1993) The attachment of minicircles to kinetoplast DNA networks during replication. **CELL** 74, 703-711
59. Perez-Morga, D.; Englund, P. (1993) The structure of replicating kinetoplast DNA networks. **J. Cell Biol.** 123, 1069-1079.
60. Maizels, N.; Weiner, A. (1988) In search for a template. **NATURE** 334, 469-470.
61. Frohman, M. (1990) RACE: Rapid Amplification of cDNA Ends. In **PCR Protocols**, Ed. Gelfand, D.; Sninsky, J.; White, T. Academic Press inc., 28.
62. Scharf, S. (1990) Cloning with PCR. In **PCR Protocols**, Ed. Gelfand, D.; Sninsky, J.; White, T. Academic Press inc, 84.
63. Vazquez, M.; Schijman, A.; Panebra, A.; Levin, M. (1992) Nucleotide Sequence of a cDNA encoding another *Trypanosoma cruzi* acidic ribosomal P2 type protein (TcP2b). **Nucleic Acids Res.**, 11, 2893.
64. D. R. Cavener (1987) Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. **Nucleic Acids Res** 15: 1353-1361.
65. Soto, M; Requena, J.; García, M; Gómez, L.; Navarete, I. And Alonso, C. (1993) Genomic organization and expression of two independent gene arrays coding for two antigenic acidic ribosomal proteins of *Leishmania*. **J. Biol. Chem.**, 268, 21835-21843.
66. B. P. Dalrymple & J. M. Peters (1992) Identification of L10e/L12e ribosomal protein genes in *Babesia bovis*. **Nucleic Acids Res** 20: 2376.
67. Vazquez, M.; Schijman, G.; Levin, M. (1992) Nucleotide sequence of a cDNA encoding a *Trypanosoma cruzi* acidic ribosomal P1 type protein. **Nucleic Acids Res.**, 10, 2599.

68. Wool, I.G., Chan, Y.L., Gluck, A. and Susuki, K. (1991) The primary structure of rat ribosomal proteins P0, P1 and P2 and a proposal for a uniform nomenclature for mammalian and yeast ribosomal proteins. **Biochimie** 73, 861-870.
69. Vazquez, M., Schijman, A., Levin, M. (1994). A short interspersed repetitive element provides a new 3' acceptor site for trans-splicing in certain ribosomal P2 β protein genes of *Trypanosoma cruzi*. **Mol. Biochem. Parasitol.** 64 327-336.
70. Schijman,AG, Vazquez, MP, Ben Dov, C, Ghio, S, Lorenzi, H, Levin, MJ. (1995). Cloning and sequence analysis of the TcP2 β cDNA variants of *Trypanosoma cruzi*. **Biochim. Biophys. Acta**, 1264, 15-18.
71. Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988) λ ZAP: a bacteriophage λ expression vector with *in vivo* excision properties. **Nucl. Acids Res.** 16, 7583-7600.
72. Alexandre, S; Guyaux, N; Murphy, H.; Coquelet, H; Pays, A.; Steiner, M; Pays, E. (1988) Putative genes of a variant specific antigen gene transcription unit in *Trypanosoma brucei*. **Mol. Cell. Biol.** 8, 2367-2378.
73. Vijayasarathy, S.; Ernest, I.; Itzhaki, J.; Sherman, D.; Mowatt, M.; Michels, P.; Clayton, C. (1990) The genes encoding fructose biphosphate aldolase in *Trypanosoma brucei* are interspersed with unrelated genes. **Nucleic Acids Res.** 18, 2967-2975.
74. Cano, M., Gruber, A., Vazquez, M., Cortes, A., Levin, M., Gonzalez, A., Degrove, W., Rondinelli, E., Zingales, B., Ramirez, J.L., Alonso, C., Requena, J.M., da Silveira, J.F. (1995) Molecular karyotype of clone CL Brener chosen for the *Trypanosoma cruzi* Genome Project. **Mol. Biochem. Parasitol.** 71, 273-278.
75. Hayashi, K. (1991) PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. **PCR Methods and Applications**, 34-38.
76. Orita, M., Suzuki, Y. Sekiya, T., Hayashi, K.(1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. **Genomics** 5, 874-879.

77. Sheffield, V., Beck, J., Kwitek, A., Sandstrom, D., Stone, E. (1993) The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. **Genomics** 16, 325-332.
78. Hasan, G; Turner, M; Cordingley, J. (1984) Complete nucleotide sequence of an unusual mobile element from *Trypanosoma brucei*. **CELL**. 37, 333-341.
79. Novak, E.; de Mello, M.; Gomes, H.; Galindo, I.; Guevara, P.; Ramirez, J.; da Silveira, F. (1993) Repetitive sequences in the ribosomal intergenic spacer of *Trypanosoma cruzi*. **Mol. Biochem. Parasitol.** ,60, 273-280.
80. Vieira de Arruda, M., Reinach, F.C., Colli, W. and Zingales, B. (1990) Sequence of the 24S α ribosomal RNA gene and characterization of a corresponding pseudogene from *Trypanosoma cruzi*. **Mol. Biochem. Parasitol.** 40, 35-42.
81. Requena, J.M., Jimenez-Ruiz, A., Soto, M., Lopez, M.C. and Alonso, C. (1992) Characterization of a highly repeated interspersed DNA sequence of *Trypanosoma cruzi*: Its potential use in diagnosis and strain classification. **Mol. Biochem. Parasitol.** 51, 271-280.
82. Lupski, J.R. and Weinstock, G.M. (1992) Short, interspersed repetitive DNA sequences in prokaryotic genomes. **J. Bacteriol.** 174, 4525-4529.
83. Coburn, C.; Otterman, K.; McNeely, T.; Turco, S. and Beverly, S. (1991) Stable DNA transfection of a wide range of trypanosomatids. **Mol. Biochem. Parasitol.** 46, 169-180.
84. Bellofatto, V.; Torres-Muñoz, J. and Cross, G. (1991) Stable transformation of *Leptomonas seymouri* by circular extrachromosomal elements. **Proc. Natl. Acad. Sci.** 88, 6711-6715.
85. Lu, H. and Buck, G. (1991) Expression of an exogenous gene in *Trypanosoma cruzi* epimastigotes. **Mol. Biochem. Parasitol.** 44, 109-114.
86. Moser, D.R., Kirchhoff, L.V. and Donelson, J.E. (1989) Detection of *Trypanosoma cruzi* by DNA amplification using the Polymerase Chain Reaction. **J. Clin. Microbiol.** 27, 1477-1482.
87. Frasch, A.; Cazzulo, J.; Aslund, L.; Pettersson, U.(1991) Comparison of genes encoding *Trypanosoma cruzi* antigens. **Parasitol. Today.** 6

88. Chumakov, I.; Le Gall, I.; Billaut, A.; Ougen, P.; Soularue, P.; Guillou, S.; Rigault, P.; Bui, H.; De Tand, M.; Barillot, E.; Abderrahim, H.; Cherif, D.; Berger, R.; Le Paslier, D. and Cohen, D. (1992). Isolation of chromosome 21-specific yeast artificial chromosome from human genome library. **Nature Genetics** 1, 222-225.
89. Munroe, D.; Haas, M.; Bric, E.; Whitton, T.; Aburatani, H.; Hunter, K.; Ward, D. and Housman, D. IRE-Bubble PCR: a rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. **Genomics** 19, 506-514.
90. Levin, M.J., Vazquez, M., Kaplan, D. and Schijman, A.G. (1993) The *Trypanosoma cruzi* ribosomal P protein family: Classification and Antigenicity. **Parasitol. Today** 9, 381-384.
91. Skeiky, Y.; Benson, D.; Parsons, M. Elkon, K.; Reed, S. (1993) *Trypanosoma cruzi* acidic ribosomal P protein gene family. Novel P proteins encoding unusual cross-reactive epitopes. **J. Immunol.**, 10, 5504-5515.
92. Puerta, C.; Martin, J.; Alonso, C.; López, M. (1994) Isolation and characterization of the gene encoding histone H2A from *Trypanosoma cruzi*. **Mol. Biochem. Parasitol.**, 64, 1-10.
93. Castro, C.; Hernandez, H.; Castaneda, M. (1981) *Trypanosoma cruzi* ribosomal RNA: internal break in the large molecular mass species and number of genes. **Mol. Biochem. Parasitol.** 2, 219-233
94. Howard, B.H. and Sakamoto, K. (1990) Alu interspersed repeats: Selfish DNA or a functional gene family. **N. Biol.** 2, 759-770.
95. Ullu, E. (1982) The human Alu family of repeated DNA sequences. **Trends Biochem.** 7, 216-219.
96. Matthews, K.; Tschudi, C. and Ullu, E. (1994) A common pyrimidine-rich motif governs trans-splicing and poliadenilation of tubulin polycistronic pre-mRNA in trypanosomes. **Genes & Development** 8, 491-501.
97. Requena, J.; Soto, M. and Alonso, C. (1993) Isolation of *Trypanosoma cruzi* specific nuclear repeated DNA sequences. **Biol Res** 26: 11-18 (1993)

88. Chumakov, I.; Le Gall, I.; Billaut, A.; Ougen, P.; Soularue, P.; Guillou, S.; Rigault, P.; Bui, H.; De Tand, M.; Barillot, E.; Abderrahim, H.; Cherif, D.; Berger, R.; Le Paslier, D. and Cohen, D. (1992). Isolation of chromosome 21-specific yeast artificial chromosome from human genome library. **Nature Genetics** 1, 222-225.
89. Munroe, D.; Haas, M.; Bric, E.; Whitton, T.; Aburatani, H.; Hunter, K.; Ward, D. and Housman, D. IRE-Bubble PCR: a rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. **Genomics** 19, 506-514.
90. Levin, M.J., Vazquez, M., Kaplan, D. and Schijman, A.G. (1993) The *Trypanosoma cruzi* ribosomal P protein family: Classification and Antigenicity. **Parasitol. Today** 9, 381-384.
91. Skeiky, Y.; Benson, D.; Parsons, M. Elkon, K.; Reed, S. (1993) *Trypanosoma cruzi* acidic ribosomal P protein gene family. Novel P proteins encoding unusual cross-reactive epitopes. **J. Immunol.**, 10, 5504-5515.
92. Puerta, C.; Martin, J.; Alonso, C.; López, M. (1994) Isolation and characterization of the gene encoding histone H2A from *Trypanosoma cruzi*. **Mol. Biochem. Parasitol.**, 64, 1-10.
93. Castro, C.; Hernandez, H.; Castaneda, M. (1981) *Trypanosoma cruzi* ribosomal RNA: internal break in the large molecular mass species and number of genes. **Mol. Biochem. Parasitol.** 2, 219-233
94. Howard, B.H. and Sakamoto, K. (1990) Alu interspersed repeats: Selfish DNA or a functional gene family. **N. Biol.** 2, 759-770.
95. Ullu, E. (1982) The human Alu family of repeated DNA sequences. **Trends Biochem.** 7, 216-219.
96. Matthews, K.; Tschudi, C. and Ullu, E. (1994) A common pyrimidine-rich motif governs trans-splicing and poliadenilation of tubulin polycistronic pre-mRNA in trypanosomes. **Genes & Development** 8, 491-501.
97. Requena, J.; Soto, M. and Alonso, C. (1993) Isolation of *Trypanosoma cruzi* specific nuclear repeated DNA sequences. **Biol Res** 26: 11-18 (1993)

98. Requena, J.; Martin, F.; Soto, M.; Lopez, M.; Alonso, C. (1994) Characterization of a short interspersed reiterated DNA sequence of *Trypanosoma cruzi* located at the 3'-end of a poly(A)+ transcript. **Gene** 146, 245-250.
99. Murphy, N.B., Pays, A., Tebabi, P., Coquelet, H., Guyaux, M., Steinert, M., and Pays, E. (1987) *Trypanosoma brucei* repeated element with unusual structural and transcriptional properties. **Mol. Biol.** 195, 855-871.
100. Kendall, G., Wilderspin, A.F., Ashall, F., Miles, M.A. and Kelly, J.M. (1990) *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase does not conform to the 'hotspot' topogenic signal model. **EMBO J.** 9, 2751-2758.
101. Liv, J.; Salinas, G.; Gajedram, N.; Muthui, D.; Muyldermans, S.; Hamers, R. (1992) Recombination associated with short direct repeats in *Leishmania mexicana*. **Mol. Biol. Parasitol.** 50, 351-354.
102. Moser, D.R., Kirchhoff, L.V. and Donelson, J.E. (1989) Detection of *Trypanosoma cruzi* by DNA amplification using the Polymerase Chain Reaction. **J. Clin. Microbiol.** 27, 1477-1482.
103. Wincker, P.; Roizes, G.; Goldemberg, S. (1990) Characterization of a *Trypanosoma cruzi* specific nuclear repeated sequence. **Mol. Biochem. Parasitol.** 41, 147-152.
104. F. Martin, C. Maranon, M. Olivares, C. Alonso and M. C. Lopez (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repair enzymes. **J. Mol. Biol.** 247, 49-59